

**"FORMATOS DE ARQUIVOS DE SOM NA INTERNET:  
UMA VISÃO CONTEMPORÂNEA – USOS,  
EXPECTATIVAS E TENDÊNCIAS"**

Marcelo José Baasch Filomeno

**"FORMATOS DE ARQUIVOS DE SOM NA INTERNET:  
UMA VISÃO CONTEMPORÂNEA – USOS,  
EXPECTATIVAS E TENDÊNCIAS"**

Dissertação apresentada ao  
Programa de Pós-Graduação em  
Engenharia de Produção da  
Universidade Federal de Santa Catarina  
como requisito parcial para a obtenção  
do título de mestre em  
Engenharia de Produção,  
(Área de Concentração: Gestão Integrada do Design)

Orientador: Prof. Milton Luiz Horn Vieira, Dr.

Florianópolis  
2003

Marcelo José Baasch Filomeno

**"FORMATOS DE ARQUIVOS DE SOM NA INTERNET:  
UMA VISÃO CONTEMPORÂNEA – USOS,  
EXPECTATIVAS E TENDÊNCIAS"**

Esta dissertação foi julgada e aprovada para a  
obtenção do título **de Mestre em Engenharia de  
Produção no Programa de Pós-Graduação em**  
Engenharia de Produção da  
Universidade Federal de Santa Catarina

Florianópolis, 18 de novembro de 2003.

---

Prof. Édson Pacheco Paladini, Dr.  
Coordenador do Curso de Pós-Graduação  
Em Engenharia de Produção

**BANCA EXAMINADORA**

---

Prof. Mílton Luiz Horn Vieira, Dr.  
**Orientador**

---

Prof. Luís Fernando Figueiredo, Dr.

---

Prof. Júlio César da Silva, Dr.

*Dedicatória*

À minha Querida Mãe (*In Memoriam*), que no meio deste mundo tão injusto e violento, me fez ver que a principal característica de um verdadeiro sábio não é a inteligência, mas a TOLERÂNCIA.

## *Agradecimentos*

Aos colegas Cláudio “The Fly”, Daniel “chatuba”, Thomas e mais todo o pessoal do laboratório da sala 117 (Projeto Geometrando) pela troca de novos conhecimentos, pelo estresse, pela alegria e pelas churrascadas.

Ao meu orientador, prof. Dr. Milton Luiz Horn Vieira,  
pela confiança, empenho e amizade.

Aos meus 4 amigos da mão esquerda:  
Sisto, João, Giuseppe e Acácio, porque eu me forjei também com um  
pouco de suas qualidades: caráter, nobreza e inteligência.

Aos meus amigos virtuais Mong e Nhonho por terem me aturado durante  
as minhas crises de raiva mantendo-se quietos e impassíveis.

À Tati, pelo carinho, afeto e pela grande quantidade de endorfina que me  
fez produzir.

Aos meus queridos irmãos: César, Maria Cristina, Lili e Ricardo, pois eles  
sempre foram mais que meus melhores amigos.

Ao menininho inseguro, introspectivo, sensível e legal, que conseguiu de  
algum jeito, carregar um homem de quase 2 metros e 100 quilos até aqui!

## SUMÁRIO

SUMÁRIO.....	V
LISTA DE FIGURAS .....	IX
LISTA DE QUADROS .....	XII
RESUMO .....	XIII
ABSTRACT .....	XIV
CAPÍTULO 1- INTRODUÇÃO .....	1
1.1- Introdução.....	1
1.2- Justificativa e Relevância da Pesquisa. ....	4
1.3- Objetivos .....	5
1.4- Procedimentos Metodológicos. ....	5
1.5- Limitações do Trabalho. ....	7
1.6- Descrição e Organização dos Capítulos .....	8
CAPÍTULO 2- CONHECENDO ASPECTOS GERAIS DO SOM ANALÓGICO E DO MUNDO DO SOM NA INTERNET. ....	10
2.1- A Física Acústica Explica o Som. ....	10
2.2- Um Breve Histórico da Tecnologia do Som: "Do Gramofone ao DVD-audio". .....	15
2.3- Compreendendo o Som Digital. ....	22
2.4- Vantagens e Desvantagens do Som Digital. ....	29
2.5- A Revolução da Tecnologia <i>Streaming</i> . ....	31
2.6- Os Algoritmos de Compactação-descompactação (CODECS). ....	33
2.7- Tecnologias Variadas de Som na Internet : Internet-fone, WEB-rádio, Áudio e Videoconferência e <i>Voice-mail</i> . ....	40
2.8- O Fenômeno MP3, o Problema da Pirataria e Direitos Autorais.....	41
2.9- A Disputa pelo Formato Padrão. ....	44
2.10- Tipos de Conexão e Vias de Transmissão. ....	44
2.11- A Importância do <i>Design</i> de Som em um <i>Website</i> . ....	46
CAPÍTULO 3- OS FORMATOS DE ARQUIVOS DE SOM ENCONTRADOS NA INTERNET .....	49
3.1- Os Principais Formatos de Arquivos de Som na Internet.....	49
3.1.1- <u>O Grupo MPEG</u> .....	50

3.1.2- <u>MPEG- 1 camada 3 ("MP3")</u> .....	54
3.1.2.1- Histórico.....	54
3.1.2.2- Características.....	56
3.1.2.3- Usos, Vantagens e Desvantagens. ....	57
3.1.3- <u>Liquid Audio</u> .....	58
3.1.3.1-Histórico.....	58
3.1.3.2- Características.....	58
3.1.3.3- Usos, Vantagens e Desvantagens. ....	58
3.1.4- <u>A2B</u> .....	59
3.1.4.1- Histórico.....	59
3.1.4.2- Características.....	59
3.1.4.3- Usos, Vantagens e Desvantagens. ....	59
3.1.5- <u>QuickTime</u> .....	59
3.1.5.1- Histórico.....	59
3.1.5.2- Características.....	60
3.1.5.3- Usos, Vantagens e Desvantagens. ....	60
3.1.6- <u>Macromedia Flash e Shockwave</u> .....	61
3.1.6.1- Histórico.....	61
3.1.6.2- Características.....	61
3.1.6.3- Usos, Vantagens e Desvantagens. ....	61
3.1.7- <u>Wave</u> .....	62
3.1.7.1- Histórico.....	62
3.1.7.2- Características.....	62
3.1.7.3- Usos, Vantagens e Desvantagens. ....	62
3.1.8- <u>AIFF</u> .....	63
3.1.8.1- Histórico.....	63
3.1.8.2- Características.....	63
3.1.8.3- Usos, Vantagens e Desvantagens. ....	63
3.1.9- <u>Sun Audio (AU)</u> .....	64
3.1.9.1- Histórico.....	64
3.1.9.2- Características.....	64
3.1.9.3- Usos, Vantagens e Desvantagens. ....	64
3.1.10- <u>RealAudio ( RealSystem G2)</u> .....	65

3.1.10.1- Histórico.....	65
3.1.10.2- Características.....	65
3.1.10.3- Usos, Vantagens e Desvantagens. ....	66
3.1.11- <u>Windows Media</u> .....	67
3.1.11.1- Histórico.....	67
3.1.11.2- Características.....	67
3.1.11.3- Usos, Vantagens e Desvantagens. ....	67
3.1.12- <u>VQF</u> .....	68
3.1.12.1- Histórico.....	68
3.1.12.2- Características.....	68
3.1.12.3- Usos, Vantagens e Desvantagens. ....	69
3.1.13- <u>Midi</u> .....	69
3.1.13.1- Histórico.....	69
3.1.13.2- Características.....	70
3.1.13.3- Usos, Vantagens e Desvantagens. ....	70
3.1.14- <u>MOD</u> .....	72
3.1.14.1- Histórico.....	72
3.1.14.2- Características.....	72
3.1.14.3- Usos, Vantagens e Desvantagens. ....	73
3.1.15- <u>Beatnik</u> .....	73
3.1.15.1- Histórico.....	73
3.1.15.2- Características.....	74
3.1.15.3- Usos, Vantagens e Desvantagens. ....	75
3.1.16- <u>Crescendo</u> .....	75
3.1.16.1- Histórico.....	75
3.1.16.2- Características.....	76
3.1.16.3- Usos, Vantagens e Desvantagens. ....	76
3.1.17- <u>Ogg Vorbis</u> .....	76
3.1.17.1- Histórico.....	76
3.1.17.2- Características.....	77
3.1.17.3- Usos, Vantagens e Desvantagens. ....	77
3.1.18- <u>QDX (Qdesign)</u> .....	78
3.1.18.1- Histórico.....	78



3.1.18.2- Características.....	78
3.1.18.3- Usos, Vantagens e Desvantagens. ....	78
3.1.19- <u>Os Formatos Menores e Menos Conhecidos</u> . ....	79
3.1.20- <u>Quadro Geral Comparativo dos Principais Formatos de Arquivos de Som em Uso Detectados na WEB</u> .....	80
<b>3.2- O Uso dos Formatos de Áudio nos Tipos de <i>Websites</i>: Empresariais- comerciais, Rádios-WEB, Institucionais e Pessoais.....</b>	<b>82</b>
3.2.1- Empresariais-comerciais .....	84
3.2.2- Rádios-WEB .....	85
3.2.3- Institucionais.....	87
3.2.4- Pessoais. ....	88
<b>3.3- Disputas, Expectativas e Tendências.....</b>	<b>91</b>
<b>CAPÍTULO 4- CONCLUSÕES.....</b>	<b>95</b>
<b>4.1- Conclusões .....</b>	<b>95</b>
<b>4.2- Recomendações para Trabalhos Futuros. ....</b>	<b>99</b>
<b>REFERÊNCIAS BIBLIOGRÁFICAS .....</b>	<b>100</b>
<b>GLOSSÁRIO .....</b>	<b>109</b>
<b>ANEXO- Lista de Arquivos de Áudio e suas Extensões .....</b>	<b>122</b>

## LISTA DE FIGURAS

Figura 1- A revolução do áudio digital está em todo lugar. Operador de uma estação de áudio digital simples gerencia as músicas (em MP3) tocadas em uma festa de casamento. ....	3
Figura 2 - Interface do buscador Google no modo diretório. ....	6
Figura 3 - Corte esquemático de uma cabeça humana na altura do canal auditivo mostrando os ouvidos externo, médio e interno. A:tímpano; B:martelo; C:bigorna; D:estribo; F:cóclea G:nervo auditivo. Em E temos os canais semicirculares que estão envolvidos com o senso de equilíbrio. ....	11
Figura 4 - Estrutura esquemática de uma onda sonora, onde temos em A=amplitude e $\lambda$ =Comprimento ....	12
Figura 5 - O fonógrafo de Edson. ....	16
Figura 6 - O grafofone de Tainter e Bell em 1885.....	16
Figura 7 - O gramofone de Berliner em 1887. ....	17
Figura 8 - Um Theremin sendo tocado. ....	18
Figura 9 -Variadas unidades de DAT: portátil (esquerda), para carro (direita) e <i>deck</i> para casa (embaixo) ao lado de fitas DAT.....	20
Figura 10 - O minidisk da Sony com dois discos. ....	21
Figura 11 -Tipos de <i>Digital Compact Cassette</i> (DCC): unidade para carro (esquerda em cima), console para casa (direita em cima) e unidade portátil (embaixo).....	22
Figura 12 - O processo de amostragem digital ( <i>codec</i> PCM) do som em A, B e C (neste exemplo amostras com 8 bits, resultando em aproximadamente 256 passos de quantificação: -127 até +127). Em C a taxa de amostragem é aumentada, ou seja, o intervalo de amostragem é diminuído. ....	24
Figura 13 - Em A,B e C o processo de quantificação visto em detalhamento crescente ( <i>codec</i> PCM). Observe que em C a linha da amostra (intervalo amostrado) não coincide com o passo de quantificação 119, resultando daí um erro: o erro de quantificação. A linha curva representa o sinal analógico. ....	26
Figura 14 - Estrutura básica interna de um arquivo de som digital. ....	27

Figura 15 - O limite teórico proposto por Nyquist: 2 pontos amostrais por ciclo ou período da onda sonora. Menos do que isso provocaria o <i>aliasing</i> . ....	27
Figura 16 - Com apenas um ponto amostral (as setas indicam os pontos amostrais) por ciclo, não podemos representar corretamente as frequências presentes na onda analógica original, ocorrendo um erro de captação de frequências denominado <i>aliasing</i> . Para captar o limite de 20 KHz audíveis aos seres humanos é que padronizou-se 44,1 KHz para a amostragem dos CD's de música (pouco mais do dobro da capacidade humana). ....	28
Figura 17 - O processo de digitalização do som em uma cena corriqueira do mundo real. ....	28
Figura 18 - Efeito de <i>clipping</i> em um sinal sonoro com amplitude alta demais para o conversor analógico-digital. ....	31
Figura 19 - O processo durante a execução de um arquivo Beatnik. ....	74
Figura 20 - Ocorrência porcentual (%) dos formatos de áudio nos <i>sites</i> amostrados do grupo empresariais-comerciais. ....	84
Figura 21 - Presença porcentual (%) dos formatos de áudio nos <i>sites</i> amostrados do grupo empresariais-comerciais que apresentavam obrigatoriamente algum tipo de áudio. Os <i>sites</i> que não possuíam áudio foram desconsiderados nesta representação. ....	85
Figura 22 - Ocorrência porcentual (%) dos formatos de áudio nos <i>sites</i> amostrados do grupo rádios-WEB. Aqui neste grupo apenas os formatos para transmissão da programação (via <i>streaming</i> ) foram considerados, qualquer outro tipo de áudio estrutural dos <i>sites</i> foi desconsiderado. ....	86
Figura 23 - Ocorrência porcentual (%) dos formatos de áudio nos <i>sites</i> amostrados do grupo institucionais. ....	87
Figura 24 - Presença porcentual (%) dos formatos de áudio nos <i>sites</i> amostrados do grupo institucionais que apresentavam obrigatoriamente algum tipo de áudio. Os <i>sites</i> que não possuíam áudio foram desconsiderados nesta representação. ....	88
Figura 25 - Ocorrência porcentual (%) dos formatos de áudio nos <i>sites</i> amostrados do grupo pessoais. ....	89

Figura 26 - Presença porcentual (%) dos formatos de áudio nos *sites* amostrados do grupo pessoais que apresentavam obrigatoriamente algum tipo de áudio. Os *sites* que não possuíam áudio foram desconsiderados nesta representação..... 90

## LISTA DE QUADROS

Quadro 01 - Altura relativa de sons comuns.....	14
Quadro 02 - Quadro comparativo das características dos formatos de áudio.....	81

## RESUMO

FILOMENO, Marcelo José Baasch. **FORMATOS DE ARQUIVOS DE SOM NA INTERNET: UMA VISÃO CONTEMPORÂNEA — USOS, EXPECTATIVAS E TENDÊNCIAS**. Florianópolis, 2003. 126f. Dissertação (Mestrado em Engenharia de Produção) – Programa de Pós-graduação em Engenharia de Produção, UFSC, 2003.

A quantidade de formatos de arquivos de som na Internet atualmente é muito grande. Chegar a um consenso de um padrão ideal é praticamente impossível, visto a grande rapidez com que mudam as tecnologias dessa área e também pela disputa mercadológica entre as grandes empresas do ramo. Na área de áudio na WEB, o aspecto mercadológico se mostrou tão importante quanto o técnico. Este trabalho visa dar uma visão geral dos principais formatos de arquivos de áudio usados hoje na Internet, tentando mostrar suas características principais, vantagens e desvantagens. Dos formatos de arquivos de áudio identificados nos determinados grupos de *sites* (empresariais-comerciais, rádios-WEB, institucionais e pessoais), os mais utilizados e promissores foram: Windows Media (WMA da empresa Microsoft); MPEG-1 layer-III (MP3 do grupo MPEG), RealAudio (Real Networks) e MPEG-2 AAC (grupo MPEG). Certos formatos mais antigos ainda perduram mais por tradição do que por capacidades técnicas (WAVE, AIFF e Sun Audio-AU), enquanto outros se mantêm por terem peculiaridades técnicas únicas, como os formatos MOD e MIDI (tipo partituras digitais). Várias empresas tentam estabelecer padrões de áudio para comércio eletrônico (Liquid Audio, QDX e Quicktime), e também para montagem de *sites* (Beatnik e Crescendo). Certos formatos devidos às suas características técnicas e de fidelidade de som se mostraram mais indicados para uma ou outra plataforma de apresentação (CD-ROM, DVD, Internet).

**Palavras-Chave:** Som Digital, Internet , Formatos de Arquivos.

## ABSTRACT

FILOMENO, Marcelo José Baasch. **FORMATOS DE ARQUIVOS DE SOM NA INTERNET: UMA VISÃO CONTEMPORÂNEA — USOS, EXPECTATIVAS E TENDÊNCIAS.** Florianópolis, 2003. 126f. Dissertação (Mestrado em Engenharia de Produção) – Programa de Pós-graduação em Engenharia de Produção, UFSC, 2003.

Nowadays, the amount of sound file formats on the Internet is very large. It is virtually impossible to make a choice about the potential best global standard of a sound format due to fast technologies changes in this area and commercial disputes among large companies. When we talk about digital sound file formats the commercial point of view is so important as the technological features. This thesis intends to provide a global view of the main digital audio formats used in the Internet nowadays, trying to show its features, advantages and drawbacks. Among the researched and detected formats at different sites (companies-commerce, WEB-radios, organizations and personal sites) the most utilized and promising were: Windows Media (WMA from Microsoft), MPEG-1 layer-III (MP3 from MPEG group), RealAudio (from Real Networks) and MPEG-2 AAC (from MPEG group). Certain older formats still remain in the business due more to the tradition than to technological features (WAVE, AIFF and Sun Audio-AU) while others are still alive because of their technical peculiarities like MOD and MIDI formats (similar to digital sheet music). Many companies try to establish digital audio standards for e-commerce (Liquid Audio, QDX and Quicktime) and for site creation (Beatnik e Crescendo). Some formats (due to its technological and quality sound features) were revealed as more indicated to certain presentation platforms (CD-ROM, DVD, Internet).

**Keywords: Digital Sound, Internet , File Formats.**

## CAPÍTULO 1- INTRODUÇÃO

### 1.1- Introdução

Apesar da visão ser o sentido mais poderoso do ser humano — praticamente 1/3 do cérebro humano destina-se ao ato de capturar e processar imagens — o sentido da audição também tem seu lugar de destaque no decorrer da evolução do homem. Diz a sabedoria popular que uma imagem vale mais que mil palavras. Não se estaria exagerando se fosse dito que um som vale pelo menos quinhentas palavras. Não é à toa que as palavras “audiência”, “auditório” e “áudio” têm a mesma origem no latim: *audire*, que significa “ouvir” (MULLIN, 2002).

Desde mesmo antes de nascer, dentro do útero materno, o ser humano é alvo incessante dos batimentos cardíacos e outros ruídos do corpo da mãe. É um engano achar que o feto não pode ouvir o que se passa lá fora, ele já reage aos sons externos mostrando desagrado ou bem-estar, conforme os variados tipos de sons que chegam ao seu sistema auditivo (MULLIN, 2002).

Os sons podem transmitir as mais variadas informações, desde informações técnicas, até puras reações emocionais, que são absorvidas das mais variadas formas conforme cada ser humano e suas personalidades e histórias diferentes. Um estrondo de trovada pode causar grande pavor ou até mesmo entusiasmo e alegria em pessoas diferentes. Contudo, há sons que chegaram até a ser memorizados geneticamente (o que chamamos instinto) como uma poderosa forma de transmitir informações. Poderíamos citar o som do guizo da cascavel, considerado como um dos sons mais aterrorizantes do mundo. Mesmo um veado com poucos dias de vida de idade, ao ouvir o chocalho da cascavel, torna-se ansioso e assustado. Em função disso, pode-se perceber o grande poder de informação, seja de uma forma consciente ou sub-liminar, que um som pode conter, tanto na sua forma mais rebuscada, como o código de sons da fala, como em outras formas bem mais simples, como um tambor tribal, o latido do cão ou o ressoar das ondas na praia.

Através de muitas maneiras e tecnologias, o homem com o passar dos tempos foi se tornando capaz de controlar e criar sons com extrema sofisticação e poder.



Instrumentos musicais com poderosos sintetizadores de som de última geração, fitas digitais, *compact disk laser* (CD), DVD-áudio e tecnologias de *software*, entre muitas outras, fazem parte desta história. A era dos computadores chegou e com ela quase todo tipo de *hardware* proprietário usava seu próprio formato de arquivo para áudio digital, mas alguns formatos tornaram-se mais genericamente aplicáveis chegando a ponto de estabelecerem padrões. Com o passar do tempo começou a haver um grande empenho para o desenvolvimento de superpadrões para variados usos.

Similar à corrida espacial, que trouxe no bojo de suas pesquisas e avanços, o desenvolvimento de várias outras tecnologias – entre elas, muitas que usamos no dia-a-dia e nem imaginamos ser tecnologia de origem espacial – a *Internet* abriu um vasto portal de outros avanços tecnológicos. Novas linguagens de programação e *softwares* para os mais variados fins (controle, análise e tráfego de informação de todos os tipos) foram desenvolvidos. De todas essas tecnologias, as que mais se sobressaem diretamente aos sentidos são aquelas envolvidas com a multimídia na Internet: vídeo, imagens, textos e sons.

Quando é feita uma referência a som na Internet, logo vem à tona a sigla MP3 (MPEG-1 *layer III*), o formato de som revolucionário que mexeu com as bases da grande rede e também do mercado fonográfico se tornando uma mania mundial (Figura-1). Contudo, o formato MP3 é apenas a ponta do *iceberg*. O som transita pela grande rede através dos mais variados formatos.

A quantidade de extensões dos arquivos de som encontrados dá uma idéia da variedade existente: .WAV, .SAM, .AIF, .VOX, .AU, .SMP, .MP1, .MP2, .ASF, .SVX, .IFF, .V8, .VOC, .PAT, .IVC, .SND, .SDS, .SFR, .DIG, .SD, Quicktime, Flash e outros mais (vide Anexo para uma lista completa das extensões de formatos de som digital). Para somar ainda mais a esta complicada resenha de extensões e formatos temos que, uma certa extensão por si só, algumas vezes não chega a definir com exatidão o tipo de arquivo. Fato bem demonstrado com a extensão “.SND”, que pode indicar formatos de arquivos de som completamente diferentes. A extensão .SND poderia indicar um arquivo de uma estação Next, um Macintosh, um PC com Windows, um Amiga e até mesmo um *videogame* Tandy.

Podemos observar que o som na Internet não é apenas música para aficionados trocarem entre si, mas toma as formas mais variadas: som ao vivo e arquivado, jornalismo, livros em áudio, arquivos históricos de som, tutoriais educacionais em

áudio, áudio incluso em jogos, efeitos de som variados utilizados como parte do *design* de páginas-WEB, WEB-rádios, *e-mail* sonoros, videoconferências e WEB-telefonía.



Figura 1- A revolução do áudio digital está em todo lugar. Nesta imagem, um operador de uma estação de áudio digital portátil gerencia as músicas (em MP3) tocadas em uma festa de casamento. O contraste se faz visível quando compara-se com outras épocas em que o operador de som lidava com muitos CD's, fitas K-7's e LP's. Hoje, em um único disco-rígido de computador pode haver milhares de músicas.

Levando-se em conta o grande mercado consumidor, o som, devido ao menor consumo de largura de banda na rede, sempre está um passo à frente do vídeo. A WEB-telefonía já é fato comum em muitos locais do mundo, contudo o vídeo ainda está em estado inicial, com suas tradicionais imagens de baixa qualidade de 4 quadros por segundo e em tamanho muito pequeno. Conexões de média velocidade como padrões ADSL e a cabo (128, 256 e 512 Kbps) já suportam bem o trânsito de voz em tempo real. Entretanto, visto a grande velocidade com que as tecnologias se desenvolvem e a obsolescência de produtos comerciais nesta área (*software* e *hardware*), a tecnologia do áudio-vídeo abrangerá um número muito maior de pessoas em curto espaço de tempo. O único problema é que esta evolução

geralmente só ocorre nos países ricos e em pequenas partes da população nos países pobres, ficando a maioria da população mundial fora destes avanços.

## 1.2- Justificativa e Relevância da Pesquisa

Devido à grande quantidade de tecnologias e formatos de arquivos de sons na Internet – onde praticamente tudo evolui com grande velocidade – procura-se neste trabalho, tentar mostrar uma visão geral do que há disponível no momento na Internet no que se refere a formatos de sons. Apesar de existir o que já se poderia chamar de alguns padrões de formatos de sons, ainda existe uma grande variedade de outros formatos.

A importância econômica da Internet no mundo atual é inquestionável. Pode-se dizer que um empresário que ainda não se estabeleceu na grande rede está perdendo tempo precioso e competitividade. A tendência atual é tornar a empresa o mais virtual possível, cortando despesas e custos, acessando o maior número de possíveis clientes no menor tempo possível. E para isso, talvez um dos melhores métodos seja o de procurar ser o mais "multimídia" possível.

Pensando nisso, procurou-se mostrar os principais formatos de sons existentes hoje na Internet, fornecendo subsídios e auxílio na compreensão de tais tecnologias. Isto poderá fazer com que empreendedores e criadores de *sites* na Internet adquiram um conhecimento em sons que os ajude a otimizar tempo, dinheiro e a própria saúde física, porque saber encontrar a metodologia correta na montagem de um *site* na WEB — seja de um empreendimento empresarial visando lucro ou apenas um mero *site* pessoal— pode ser extremamente desgastante. Daí, uma visão geral atual das tecnologias de som na WEB ser de grande valia, pois assim aprimora-se em uma das pilstras básicas da informação multimídia (som-imagem-texto-vídeo-animação).

Além da parte do uso comercial e pedagógico dos tipos de tecnologias de sons na Internet, é importante ressaltar a grande relevância social, pois deficientes visuais e físicos dos mais variados níveis de dificuldades poderão se beneficiar de tais tecnologias. Entre elas poderiam ser citadas os livros falados (*e-books*), *softwares* leitores de texto por voz sintética e reconhecimento de voz.

### 1.3- Objetivos

O objetivo principal deste trabalho é o de identificar e comparar os formatos de arquivos de sons mais utilizados hoje na *World Wide Web* (Internet), procurando estabelecer suas vantagens e desvantagens. Em nível complementar procurar-se-á determinar o uso mais indicado para cada formato, procurando-se ter uma visão global das tecnologias de áudio atualmente na Internet.

### 1.4- Procedimentos Metodológicos

Em uma primeira etapa, será feita a pesquisa bibliográfica, sendo que esta será embasada tanto em livros, revistas e outros impressos, tanto quanto na Internet, pois o assunto em questão refere-se a tecnologia atual e que sofre avanços com grande rapidez.

Em uma segunda etapa, navegar-se-á pela Internet visando a análise do sistema de som de vários tipos de *sites*. Para melhor compreensão, estes *sites* serão divididos em 4 grandes grupos representativos de níveis sócio-econômicos :

- a) **Empresariais-comerciais**- incluindo indústria ,comércio, bancos e qualquer associação com fim lucrativo .
- b) **Rádios-WEB**- apesar de também poderem ser incluídas no grupo “empresariais-comerciais”, as rádios-WEB são os *sites* mais característicos do uso de som na Internet, por isso foram separados em um grupo único.
- c) **Institucionais**- ONG's, escolas, associações, clubes e *sites* governamentais (*sites* sem fins lucrativos).
- d) **Pessoais**- *sites* de criação pessoal.

Serão analisados *sites* para cada um dos grandes grupos. Com isso, procurar-se-á observar se existem padrões de formatos de arquivos de som diferenciados para cada um dos referidos grupos de *sites*. Estes *sites* serão obtidos a partir do diretório

específico do buscador *Google*. Os *sites* empresariais-comerciais serão obtidos no diretório indústrias-empresas do buscador, os *sites* pessoais no diretório de pessoais, e assim por diante. Podemos observar a interface do buscador *Google* no modo diretório na Figura-2.



Figura 2 - Interface do buscador Google no modo diretório.

Fonte: [www.google.com.br](http://www.google.com.br)

Para cada um dos principais formatos analisados procurar-se-á identificar características-chave como: criador do formato, dados de proteção de direitos autorais, capacidade de *streaming*, tipo de som (real, sintético ou híbrido), taxas de bits compatíveis, área de aplicação no mercado, vantagens, desvantagens e formas variantes.

Devido ao estudo deste trabalho exigir constante uso de termos bem técnicos, será disponibilizado em anexo desta dissertação um glossário com os termos utilizados nos capítulos.

Neste trabalho será dada preferência de uso para os termos em inglês, para facilitar o entendimento. Grande parte destes termos não tem a tradução em português ou sequer foram aportuguesados. Mesmo termos que já têm o aportuguesamento definido, como *site* (= sítio em português), foram evitados, pois estes tendem a soar estranhamente (sítio, além de local, também é a pequena propriedade onde se plantam frutas e criam-se animais).

## 1.5- Limitações do Trabalho

Para a procura e análise de formatos de áudio, será considerada apenas a *homepage* de cada *site*, pois a navegação por todo o *site*, com toda a estrutura e *links* das páginas constituintes (muitos *links* externos ao *sites*), seria de uma complexidade impraticável para um trabalho neste nível.

Este trabalho refere-se basicamente a formatos de arquivos de áudio e não a *codecs*. Certos formatos diferentes podem utilizar o mesmo *codec*. Mesmo assim, procurou-se dar também uma visão geral de alguns *codecs*. Muitas vezes a separação entre formato e *codec* é de difícil conceituação, como no caso dos formatos MP3 e QDX, pois estes são formatos de *codec* específico. Por outro lado, em formatos como o WAVE da Microsoft podem ser utilizados vários *codecs*.

Procurar-se-á ao máximo analisar a qualidade e veracidade das informações contidas nos *sites* visitados, pois é de praxe ter maiores ressalvas com o material da Internet do que com o material impresso (livros, revistas, compêndios, relatórios e outros trabalhos). Mesmo não sendo na sua grande parte muito atual, a parte bibliográfica em livros disponíveis também será pesquisada.

Não se pretende em hipótese alguma estruturar reais e certas estatísticas sobre os usos dos arquivos de áudio nos grandes grupos de *sites* da WEB (descritos nos procedimentos metodológicos), mas sim, tentar ter uma certa noção do que está presente atualmente na Internet. O tamanho da rede, a quantidade de *sites*, o caráter

mutante e o crescimento muito acelerado da Internet dificulta um estudo amostral mais profundo e com maior exatidão.

Além disso, poderia-se facilmente dispersar a atenção em formatos digitais não-WEB como os formatos digitais de som para vídeo ou cinema (Dolby Digital, DTS e Sony), além de vários outros formatos de arquivos de *softwares* de edição de áudio e formatos de áudio otimizados para jogos de computador (ver ANEXO). Em muitos momentos existem confluências entre estes formatos não-WEB e os formatos WEB. Tenta-se neste trabalho, mostrar com racionalidade e objetividade, os formatos que realmente importam atualmente ou que tiveram grande importância no passado, evitando-se um excesso de dados.

A velocidade de evolução de tecnologias na área da Internet é muito grande, e no que se refere ao áudio também não é diferente. Obviamente a probabilidade de existir algum formato mais inovador de WEB-áudio que não seja citado neste trabalho não é pequena.

Outro problema existente é a veracidade das informações. Observações anteriores a este trabalho mostram que há uma tendência de um desenvolvedor de certo formato de áudio-WEB de valorizar muito o seu formato em detrimento de outro, mostrando que existe uma disputa de *marketing* por detrás das especificações técnicas de cada formato.

## **1.6- Descrição e Organização dos Capítulos**

Da forma como é apresentada, esta dissertação se divide em 4 capítulos:

### **Capítulo 1: Introdução**

No capítulo 1 apresenta-se a introdução geral ao tema, a justificativa e a relevância da pesquisa, os objetivos básicos do trabalho e a descrição da metodologia adotada.

**Capítulo 2:** Conhecendo aspectos gerais do som analógico e do mundo do som na Internet

O capítulo 2 consiste da revisão de literatura, tratando dos principais conceitos e aspectos do som como fenômeno físico, a evolução histórica das tecnologias de som, a compreensão de características gerais do som digital e aspectos atuais do som na Internet.

**Capítulo 3:** Os formatos de arquivos de som encontrados na Internet

No capítulo 3 apresenta-se a descrição dos principais formatos de áudio encontrados atualmente na Internet, junto com o resultado da pesquisa da ocorrência dos vários formatos em grupos de *sítes* específicos e posteriormente também os aspectos da evolução, expectativas e tendências destes formatos.

**Capítulo 4:** Conclusões

No capítulo 4 são apresentadas as conclusões gerais sobre o trabalho, as limitações necessárias empregadas e as recomendações para o aprimoramento e desenvolvimento de trabalhos futuros em tópicos importantes, mais específicos e aprofundados referentes ao tema em questão.



## **CAPÍTULO 2- CONHECENDO ASPECTOS GERAIS DO SOM ANALÓGICO E DO MUNDO DO SOM NA INTERNET**

### **2.1- A Física Acústica Explica o Som**

O ser humano consegue captar e ouvir o som graças ao seu sistema auditivo que é dividido em três partes: o ouvido externo, o médio e o interno. O ouvido externo é constituído pelo pavilhão auditivo (orelha) e pelo canal auditivo. Aqui o som é captado pelo pavilhão que funciona como uma antena receptora e é enviado pelo canal auditivo até o tímpano (início do ouvido médio), então com a vibração deste, a energia sonora é transformada em vibração mecânica, pois põe a vibrar em seqüência os menores ossos do corpo: o martelo, a bigorna e o estribo. Por último, o estribo ao vibrar, produz uma onda de compressão dentro da cóclea (ouvido interno) que é cheia de um líquido, sendo que em suas paredes existem cílios que são ativados pelo movimento do líquido produzindo um campo elétrico que é transmitido ao cérebro através do nervo auditivo (Figura-3) (VILELA, 2003).

O som não é nada mais do que uma onda mecânica, uma sucessão de compressões e expansões das moléculas de gases que compõem o ar, que se altera no tempo e espaço. Esta onda sonora é caracterizada por ter um período, que é o tempo decorrido em um fenômeno periódico para a passagem do sistema em dois estados idênticos formando assim um ciclo e completando um comprimento de onda. O número de ciclos (pode-se também dizer períodos) por segundo é o que determina a frequência do som, que é medida em Hz (Hertz). Assim, quando diz-se que um som tem 50 Hz é porque ocorre 50 ciclos desta onda em 1 segundo. O outro valor que determina uma onda sonora é a sua amplitude, que nada mais é do que a intensidade de som, a força de deslocamento (pressão) das moléculas de ar, sendo medida em decibéis (dB) que é uma escala logarítmica (Figura-4) (DOBRIAN, 2001).

A velocidade de uma onda sonora depende do meio em que ela se propaga. Devido a duas características principais da matéria que afetam a velocidade do som (propriedade de inércia: densidade da matéria e propriedade elástica), temos que o som se propaga mais rapidamente no sólido do que no líquido e no meio líquido mais

rápido do que no gasoso. A velocidade do som ao nível do mar e 20 °C é de aproximadamente 343 m/s (THE PHYSICS CLASSROOM, 2002a).

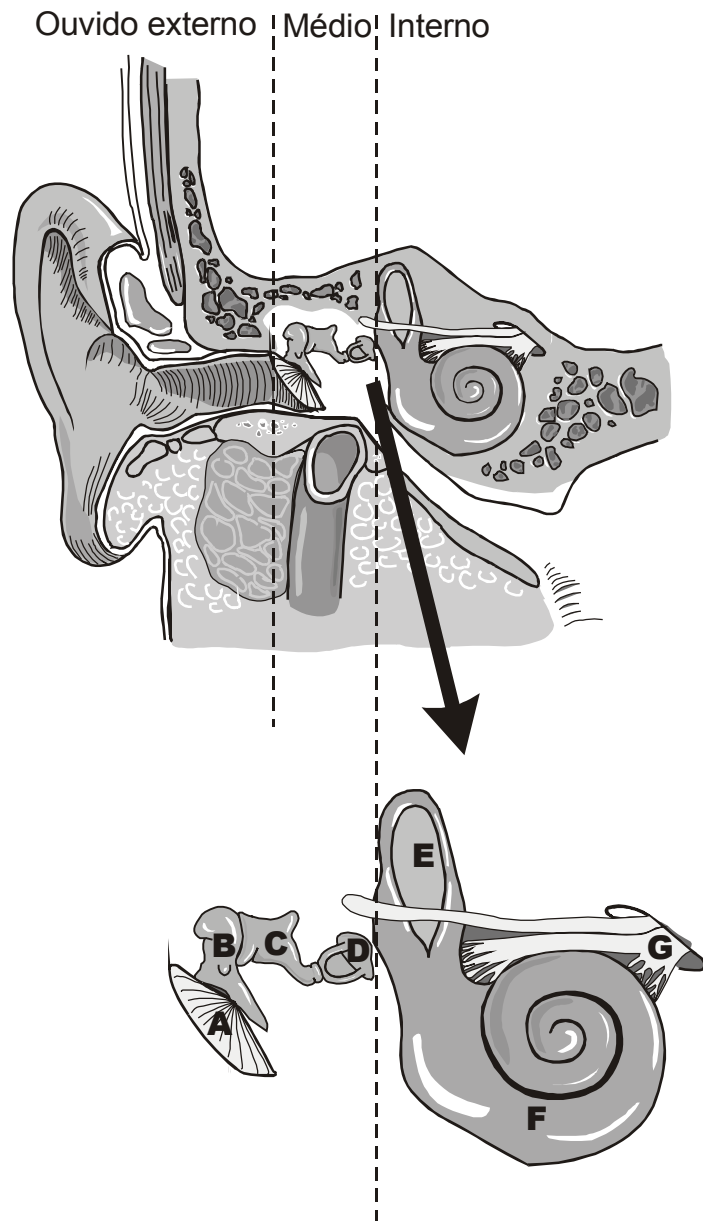


Figura 3 - Corte esquemático de uma cabeça humana na altura do canal auditivo mostrando os ouvidos externo, médio e interno. A:tímpano; B:martelo; C:bigorna; D:estribo; F:cóclea G:nervo auditivo. Em E temos os canais semicirculares que estão envolvidos com o senso de equilíbrio.

Fonte: Clipart adaptado da empresa Corel.

O ser humano raramente ouve diretamente uma onda sonora, ela antes é mascarada ou alterada por milhares de pequenas reflexões devido a características físicas do ambiente. O som também é alterado pelo material ou substância na qual através ele viaja ( BEGGS & THEDE, 2001).

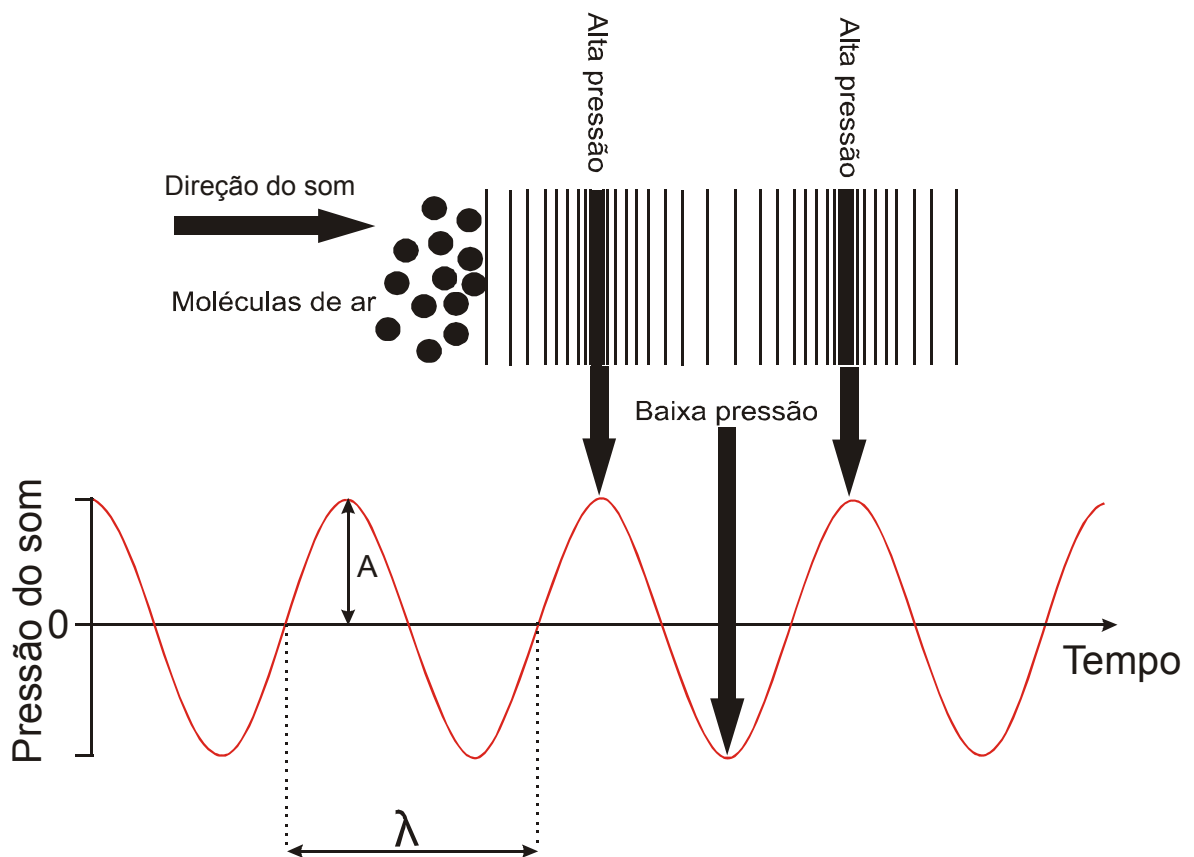


Figura 4 - Estrutura esquemática de uma onda sonora, onde temos em  $A$ =amplitude e  $\lambda$  =Comprimento

Fonte: modificado de (<http://www.physicsclassroom.com/Class/sound/U11L1c.html>).

Na realidade, durante a propagação do som observam-se os fenômenos gerais da propagação ondulatória. Dada sua natureza longitudinal, o som não pode ser polarizado como a luz; sofre, entretanto, os demais fenômenos, a saber (FEIRA DE CIÊNCIAS, 2002):

- a) difração (propriedade que a onda apresenta em contornar os obstáculos que encontra durante sua propagação);
- b) reflexão (obedece às leis da reflexão ondulatória nos meios materiais elásticos e suas conseqüências, visto que o ouvido humano discerne duas excitações breves e sucessivas apenas se o intervalo de tempo que as separa é maior ou igual a **1/10** do segundo );
- c) refração (fenômeno que caracteriza o desvio sofrido pela frente da onda sonora, que geralmente ocorre, quando ela passa de um campo ondulatório – por exemplo, ar – a outro de compressibilidade diferente – por exemplo, água);
- d) interferência (é a conseqüência da superposição de ondas sonoras);

**e) efeito Doppler** (é a consequência do movimento relativo entre o observador e a fonte sonora, o que determina uma modificação aparente na frequência do som percebida pelo observador).

Através de longas pesquisas determinou-se que a faixa de frequências audíveis para os seres humanos vai de 20 até 20.000 Hz. Os sons abaixo de 20 Hz são denominados infra-sons e acima de 20.000 Hz são denominados de ultra-sons. Os humanos não estão sozinhos na capacidade de detectar uma grande faixa de frequências. Os cães podem detectar frequências tão baixas quanto 50 Hz e tão altas quanto 45.000 Hz. Os gatos vão de 45 a 85.000 Hz. No ponto extremo de altas frequências temos os animais que se utilizam de eco-localização para navegação e alimentação, entre estes, citamos os morcegos (até 120.000 Hz) e golfinhos (200.000 Hz). Os elefantes estão no extremo de baixas frequências com capacidade de detectar até 5 Hz (THE PHYSICS CLASSROOM, 2002b).

Para cada década de idade que envelhecemos, o nosso limite superior audível, que é de cerca de 20 KHz quando nascemos, cai cerca de 2 KHz por vez. Contudo, isto varia de pessoa para pessoa. O limite inferior (cerca de 20Hz) é mais constante (HARARD; PAULISSEN, 1992).

O ser humano é mais sensível a frequências mais baixas (na faixa de 300Hz até 7.000 Hz) do que em outras frequências e também a baixas amplitudes. Isto provavelmente é um mecanismo evolucionário que demonstra a importância de se ouvir a voz humana normal que está inclusa nesta faixa, além de outros sons da natureza, como ruídos de predadores (DOBRIAN, 2001).

Quando um tom de som é tão complexo que ele contém muitas diferentes frequências sem relação matemática alguma, nós percebemos o som como ruído. Um som com muitas frequências e amplitudes aleatórias (todas as frequências presentes em igual proporção) é um som semelhante à estática conhecido como ruído branco (definição análoga à da luz branca que contém todas as frequências do espectro visível). A combinação das frequências e suas amplitudes que estão presentes em um som é chamado de espectro. Cada uma das frequências que compõe um tom complexo é chamada de frequência parcial. Quando as parciais (frequências componentes) em um tom complexo são todas múltiplas inteiras da mesma frequência fundamental, o som é denominado espectro harmônico. Cada componente do espectro harmônico é chamado de harmônico parcial ou simplesmente harmônico.

A soma destas frequências harmônicas ainda resulta em uma onda periódica tendo a frequência fundamental. Desta forma, as frequências harmônicas componentes se fundem em um tom bem definido e homogêneo (DOBRIAN, 2001).

Na prática, a intensidade das mudanças na pressão do ar deve ser maior que cerca de um bilionésimo da pressão atmosférica, mas não maior que cerca de um milésimo da pressão atmosférica. Destes cálculos podemos ver que o som mais suave que podemos ouvir tem cerca de um bilionésimo da intensidade do som mais alto que podemos lidar sem sofrer qualquer lesão da audição. Nosso subjetivo senso de “altura do som” não é diretamente proporcional à amplitude (intensidade). Calculamos a percepção da altura de um som através das amplitudes relativas entre dois sons. Sendo assim, como o som mais suave que podemos ouvir tem cerca de um bilionésimo da intensidade do som mais alto que podemos lidar sem lesão, temos uma escala plausível (faixa de audição sem lesão) que vai de 0 a 1 bilhão para a sensibilidade a diferentes amplitudes (DOBRIAN, 2001). No Quadro-1 pode-se observar exemplos corriqueiros dessas faixas de sons.

Quadro 1- Altura relativa de sons comuns.

Nível relativo	NPS (dB)	Tipo de som
10.000.000X	<b>140</b>	Um tiro com Colt 45 (7m)
	<b>130</b>	Limiar da dor
1.000.000X	<b>120</b>	Decolagem de avião a jato (60 m)
	<b>110</b>	Concerto de Rock (3m)
100.000X	<b>100</b>	Música clássica alta
	<b>90</b>	Tráfego pesado de rua (1,5 m de distância)
10.000X	<b>80</b>	Cabina de avião a jato de carreira
	<b>70</b>	Conversação comum (1m de distância)
1.000X	<b>60</b>	
	<b>50</b>	Casa suburbana comum à noite
100X	<b>40</b>	Auditório quieto
	<b>30</b>	
10X	<b>20</b>	Sussuro(1,5m)
	<b>10</b>	Folhas ao vento
Nível de referência	<b>0</b>	Limiar do audível

Fonte: (FRIES, 2000)

Devido ao tamanho da escala e ao modo de percepção da audição humana, usa-se uma escala logarítmica para se comparar amplitudes (o ouvido humano é menos sensível a altas amplitudes). O nosso senso de altura do som varia consideravelmente dependendo da frequência do som sendo considerado e também da idade do ouvinte. Pode-se dizer que a escala de amplitude varia de 0 a 1 bilhão e a escala de altura vai de 0 a alguns milhares na faixa de intensidades sem lesão à audição (THE PHYSICS CLASSROOM, 2002c).

O limiar da dor é de cerca de 120 decibéis (dB), sendo que a altura do som percebida diminui cerca de 6 dB cada vez que você duplica sua distância da fonte de som (LINDSTROM, 1996).

## **2.2- Um Breve Histórico da Tecnologia do Som: "Do Gramofone ao DVD-audio"**

A primeira tentativa de gravar o som em um meio físico foi feita pelo francês Leon Scott, através de seu fononautógrafo, basicamente um precursor dos fonógrafos de cilindro. Contudo, o que Leon conseguiu foi apenas gravar as vibrações da voz e não a voz em si. Tratava-se aqui mais de uma curiosidade do que um dispositivo útil. Charles Cros, um professor e poeta, projetou no início de 1877 as bases do fonógrafo a disco, entretanto, não teve dinheiro para registrar e implementar a idéia. A história da gravação do som poderia ter sido bem diferente se Cros tivesse condições financeiras, pois os seus fundamentos estavam corretos (MENLO PARK EDISON MUSEUM, 2002).

Quando Thomas Edson em 1877 gravou em seu fonógrafo de cilindro metálico (Figura-5) movido à manivela as palavras "*Mary had a little lamb*" ("Mary tinha um carneirinho"), provavelmente mal imaginava a grande revolução que iniciara. Pela primeira vez a voz humana tinha sido registrada em um meio físico.

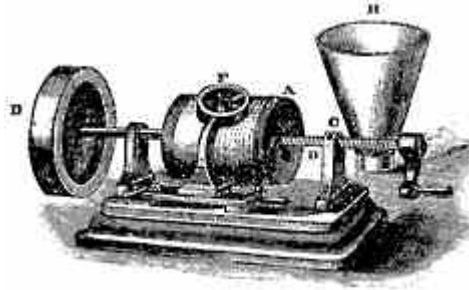


Figura 5 - O fonógrafo de Edson.

Fonte: (<http://www.snaithprimary.eril.net/edison.htm>)

O que se seguiu logo depois, foi o aparecimento de novos fonógrafos: o grafofone criado por Chichester Bell e Charles Tainter em 1885 (Figura-6) e o gramofone em 1887 criado por Emile Berliner (Figura-7), onde este já usava um disco para gravação e não mais um cilindro (SCHOENHERR, 2001).



Figura 6 - O grafofone de Tainter e Bell em 1885.

Fonte: (<http://history.sandiego.edu/gen/recording/tainter09.html>)

Em 1888 o gramofone foi melhorado e as primeiras cópias em massa de uma gravação de mídia de som haviam sido feitas através de discos de borracha dura vulcanizada, prensadas de um disco *master* em zinco (SCHOENHERR, 2001).

Com o passar dos anos, os fonógrafos, tanto cilíndricos quanto em disco, foram sendo aprimorados juntos com o tipo de mídia. Motores elétricos foram incorporados. Os cilindros tiveram o material de cobertura mudados algumas vezes (vários tipos de metais), e as agulhas de captação chegaram às de diamante (SCHOENHERR, 2001).

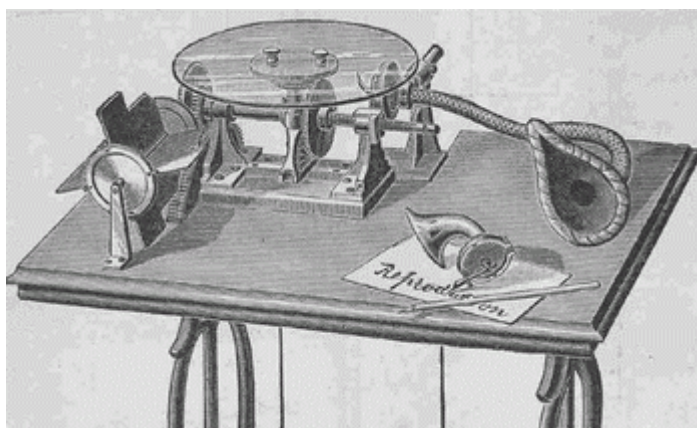


Figura 7 - O gramofone de Berliner em 1887.

Fonte: (<http://history.acusd.edu/gen/recording/images4/14120101-019b.jpg>)

Por volta de 1913 Thomas Edson, com seu fonógrafo cilíndrico, tinha se rendido a nova tecnologia dos fonógrafos com discos feitos em plástico similar ao *bakelite* (o primeiro plástico artificial patenteado na história) (SCHOENHERR, 2001).

O primeiro instrumento eletrônico que apareceu, foi desenvolvido pelo inventor russo Leon Theremin em 1919. O Theremin, como acabou sendo chamado a invenção, era um estranho instrumento manual, que se tocava ao acionar duas varas de metal ligadas a um grande estojo de madeira (Figura-8). O movimento dos braços controlava as frequências e o volume do instrumento. Ele produzia os tradicionais sons de discos voadores marcianos dos filmes dos anos 1950 e 1960 (PERSONAL COMPUTER WORLD, 2001b). Apesar de tão antigo, o Theremin tem ressurgido com vários músicos modernos, entre eles, Rita Lee e Pato Fu no Brasil e várias outras bandas em nível mundial.

Por volta de 1931, os discos de vinil já começavam a invadir o mercado de música. Os discos profissionais de gravação utilizados nas rádios nesta época eram os de cobertura de acetato de celulose e permaneceram como o meio padrão de gravação até que a fita magnética fosse adotada em 1948. Também em 1931 Alan Blumlein patenteou o método de gravação em estéreo na Inglaterra (SCHOENHERR, 2001).

Os primórdios da fita magnética remontam a 1928, porém, só em 1936, é que as empresas BASF e TELEFUNKEN fizeram a primeira gravação ao vivo de um concerto musical com um gravador de fita magnética, o magnetofone, baseado na tecnologia do Dr. Fritz Pfleumer patenteada na Alemanha (SCHOENHERR, 2001).





Figura 8 - Um Theremin sendo tocado.

Fonte: (<http://www.termenvoks.com/theremin.files/image002.jpg>)

O conceito de som digital surge apenas em 1937 com o desenvolvimento do PCM (*Pulse Code Modulation*) por Alec H. Reeves na França (ver item 2.3). Era então o melhor método para armazenar áudio num formato digital. No entanto, a tecnologia só foi posta em prática em 1962, quando o desenvolvimento dos circuitos integrados tornou possível com que a codificação PCM fosse usada para transmissões por telefone (PERSONAL COMPUTER WORLD, 2001b).

Em 1945 os americanos se apoderaram da tecnologia de gravação magnética do magnetofone (que usava fita BASF de 6.5 mm com cerca de 1.000 metros coberta com partículas de ferro e com duração de cerca de 20 minutos). Logo em 1948, os americanos lançariam a sua versão de gravador de fitas. Fitas estas de acetato com cobertura de óxido de ferro desenvolvidas pela "*Minnesota Mining and Manufacturing*", a conhecida **3M** (SCHOENHERR, 2001).

Por volta de 1949 já havia alguns padrões de discos (velocidades de 78, 45 e 33-1/3 RPM), agora já com 2 lados, feitos de vinil ou poliestireno. Em 1958 o primeiro LP (33-1/3 RPM) estéreo foi vendido, e assim um padrão mundial foi adotado para discos (SCHOENHERR, 2001).

Nas décadas de 1950, 1960 e 1970, o que se seguiu foi um avanço maior da gravação de vídeo com o som incluso, do que o som puro propriamente dito. Dentre as datas marcantes deste período poderíamos destacar (SCHOENHERR, 2001):

- em 1975: surgimento do padrão betamax desenvolvido pela Sony;
- em 1976: surgimento do padrão VHS desenvolvido pela JVC;
- em 1978: a Pioneer cria o LaserDisk.

Em 1972 foi criado o primeiro gravador PCM que utilizava cassetes de vídeo como meio de armazenamento e tinha um alcance de 87dB. Em 1976 foi realizada a primeira gravação de 16 bits, através de um gravador de cassetes digital feito à mão (intitulado *Soundstream* e criado por Thomas G. Stockham) (PERSONAL COMPUTER WORLD, 2001b).

A década de 1980 voltou a acelerar o passo nas tecnologias dedicadas exclusivamente ao som. No início desta década, a Sony lança o criativo produto "Walkman", um *player* de fita K-7 portátil que acabou mostrando ao mundo a grande necessidade que as pessoas tinham do som portátil.

Em 1982 os primeiros CD's (5 polegadas) são vendidos, e a ponte entre a indústria do som com a revolução digital dos computadores está formada (o som analógico finalmente encontra um método de digitalização de massa: ver item 2.3). Contudo, nesta época, os equipamentos para a amostragem digital do som eram bem inferiores ao de hoje, fazendo com que muitos resistissem com afinco à adoção dos CD's, dizendo que os mesmos tinham qualidade de som inferior ao LP de vinil (SCHOENHERR, 2001).

Por volta de 1983, o número de áudio-cassetes já superava o de LP's em vendas, este era o início da derrocada dos discos em vinil. Contudo durante a década de 1990 o final do reinado do áudio-cassete também ocorreria (SCHOENHERR, 2001).

Em 1985 Sony e Philips lançaram o padrão CD-ROM, disco com leitura a laser para computadores com tecnologia igual ao do áudio-CD. Já em 1987, os primeiros *players* de fitas digitais (DAT-*digital audio tape*) são lançados no mercado (Figura-9) e depois o minidisk (Sony) (Figura-10) aparece em 1992 (SCHOENHERR, 2001). Baseado na tecnologia de fita DAT, é desenvolvida a tecnologia de fita DDS para computadores. Este último tipo de fita é erroneamente chamada de DAT por muitas pessoas, mas, nem são compatíveis.



Figura 9 -Variadas unidades de DAT: portátil (esquerda), para carro (direita) e *deck* para casa (embaixo) ao lado de fitas DAT.

Fonte: (<http://perso.club-internet.fr/farzeno/edat.htm>)

Também nos anos 1990 foi lançado o *Digital Compact Cassette* (Figura-11) que acabou por falhar e não causar boa impressão no mercado consumidor (DISLEY, 2001).

Uma grande revolução ocorreu em 1990 quando a Sony e a Philips produziram o padrão para o CD-R (CD gravável). Os regraváveis (CD-RW) surgiram apenas em 1997. Repentinamente, os apreciadores de músicas podiam fazer cópias de seus CD's digitalmente sem perda e a indústria fonográfica começava a ficar atenta à esta questão, o que daria origem aos primeiros tratados de proteção contra cópias e direitos autorais (SMARTCOMPUTING, 2001).

Logo em 1997, foi lançado o DVD que tinha capacidade variável entre 4,7GB a 17,1 GB (conforme o número de lados e camadas) e tinha a dimensão e aparência igual a um CD-R comum. Apenas em 2000, é que foi lançado o outro padrão de DVD, o DVD-áudio, sendo específico para o mercado de som e projetado para suportar a melhor qualidade de som possível usando o *codec* PCM (sem compressão) ou uma compactação sem perda: o *Meridian Lossless Packing* (MLP). O DVD de vídeo, em geral, usa *codecs* com perda como o MPEG-1, o MPEG-2, o DTS, o Dolby AC-3 ou

até mesmo o PCM (não-comprimido) para a codificação de som (DISCTRONICS, 2003).



Figura 10 - O minidisk da Sony com dois discos.

Fonte: (<http://web.media.mit.edu/~bnhubert/dossier/expeditions/belize/images/digital-md-recorder.jpg>)

Em 1999 Sony e Philips lançam uma tecnologia para competir com o DVD-áudio: o SACD (Super Audio CD). É um disco baseado no CD normal de áudio, contudo usando um sistema de codificação de 1 bit com taxas de amostragens 64 vezes maior que o CD normal. Apresenta som multicanal, sistema contra cópias melhorado e espaço para som, vídeo e texto.

Mesmo antes do aparecimento dos SACD's já existiam CD's de música com capacidade interativa de um CD-ROM com *software*, eram os CD's *plus*.

Temos assim que o som digital percorreu um longo caminho num curto espaço de tempo. Qualquer computador pessoal mais atualizado pode codificar áudio para MP3, WMA, RA e outros *codecs* num instante, agir como um pequeno estúdio de som, lidar com reconhecimento de voz e criar sons de síntese em tempo real. Mas ainda há muito espaço para o desenvolvimento. A história do desenvolvimento dos *codecs* e tipos de mídia relacionados ao som digital acabam por se entrecruzar tornando-se quase interdependentes, ou seja, para certos tipos de mídias, sistemas de codificação diferentes são utilizados.



Figura 11 -Tipos de *Digital Compact Cassette* (DCC): unidade para carro (esquerda em cima), console para casa (direita em cima) e unidade portátil (embaixo).

Fonte: (<http://home.swipnet.se/~w-59099/dcc/Dcc-mode.htm>).

## 2.3- Compreendendo o Som Digital

As variações de amplitude e frequência de uma onda sonora ao serem captadas por um microfone (que é um sensor que se baseia no sentido da audição humana) são transformadas em sinais elétricos que variam continuamente com o tempo, ou seja, as pequenas variações de voltagem do sinal de som analógico é uma função do tempo:

$$S = f(t)$$

Para se transformar variações contínuas de voltagens em valores discretos digitais ou binários, tem-se que seguir os seguintes passos:

**Amostragem** (Figura-12): os valores do sinal analógico são amostrados em intervalos temporais de periodicidade constante. O valor amostrado oriundo do sinal analógico é mantido constante até o próximo intervalo de amostragem. Segundo o Teorema de Nyquist, se um sinal analógico tem componentes de frequência de até  $x$  Hz, a taxa de amostragem (para digitalizar este sinal analógico) deve ser ao menos

de 2x Hz, ou seja, o dobro. Na prática, em geral, vai um pouco acima do dobro. Foi daí que obteve-se o padrão de 44,1 KHz do CD-áudio, pois esta taxa de amostragem cobriria as freqüências dos sons da faixa audível para os seres humanos (20 Hz -20 KHz). Pelo mesmo motivo a voz humana é digitalizada com uma taxa de amostragem de 8 KHz, pois os componentes principais de freqüência da voz humana estão dentro de 3,1 KHz. Para evitar conceitos errôneos deve-se ter em mente uma coisa: a freqüência do som é uma característica do som em si, enquanto que freqüência ou taxa de amostragem (que é dado em amostras por segundo) é um parâmetro do processo de digitalização. São parâmetros completamente diferentes, apesar de ambos serem medidos em **hertz** (WILLRICH, 2000).

**Quantificação** (Figura-13)- transformar amostras contínuas em valores discretos chama-se quantificação. Na técnica mais padrão de quantificação, a modulação PCM (*Pulse Coded Modulation*), dividiu-se o domínio da amplitude do sinal em um número fixo de intervalos. Cada intervalo tem um mesmo tamanho e recebe um número. Na Figura-13 estes intervalos são numerados de 0 a 127 e 0 a -127 (tendo-se como exemplo uma amostragem de 8 bits de resolução). Assim, temos que cada amostra que cair dentro de um intervalo será atribuído o valor daquele intervalo. O tamanho do intervalo de quantificação é chamado de **passo de quantificação**. É justamente por este fato de tornar um intervalo de valores contínuos em um único e discreto valor (através da quantificação), e da manutenção deste mesmo valor durante todo o intervalo amostral (que é dado pela taxa de amostragem), que se consegue o caráter de confiabilidade e tolerância a erros e ruídos no sinal sonoro digital em relação ao analógico. O sinal digital "é ou não é", porém, no sinal analógico sempre temos o "mais ou menos". Contudo, o erro de quantificação, que é devido ao erro de aproximação e arredondamento que ocorre entre o valor amostrado e o valor inteiro permitido que pode ser registrado no passo de quantificação (Figura-13c), acaba introduzindo um ruído aleatório em cada amostra (*sample*), conhecido como ruído de quantificação. Este erro é inevitável mas pode ser reduzido a um nível aceitável através do aumento da resolução da amostra. Este ruído não está presente na amostra original, ele é adicionado ao sinal digital pela natureza imprecisa da quantificação. A razão da amplitude máxima do sinal com o erro máximo de quantificação é chamado de SQNR (*signal-to-quantization-noise-ratio*). Como uma

regra prática estabelece-se que cada bit a mais na resolução de uma amostra adiciona-se 6dB (decibéis, medida comumente usada para este fim) à SQNR. Sendo assim, um som amostrado com 16 bits de resolução tem uma SQNR de 96 dB (16 bits X 6dB =96). Em resumo, quanto maior a resolução da amostra menor o ruído e melhor o sinal (DOBRIAN, 2001).

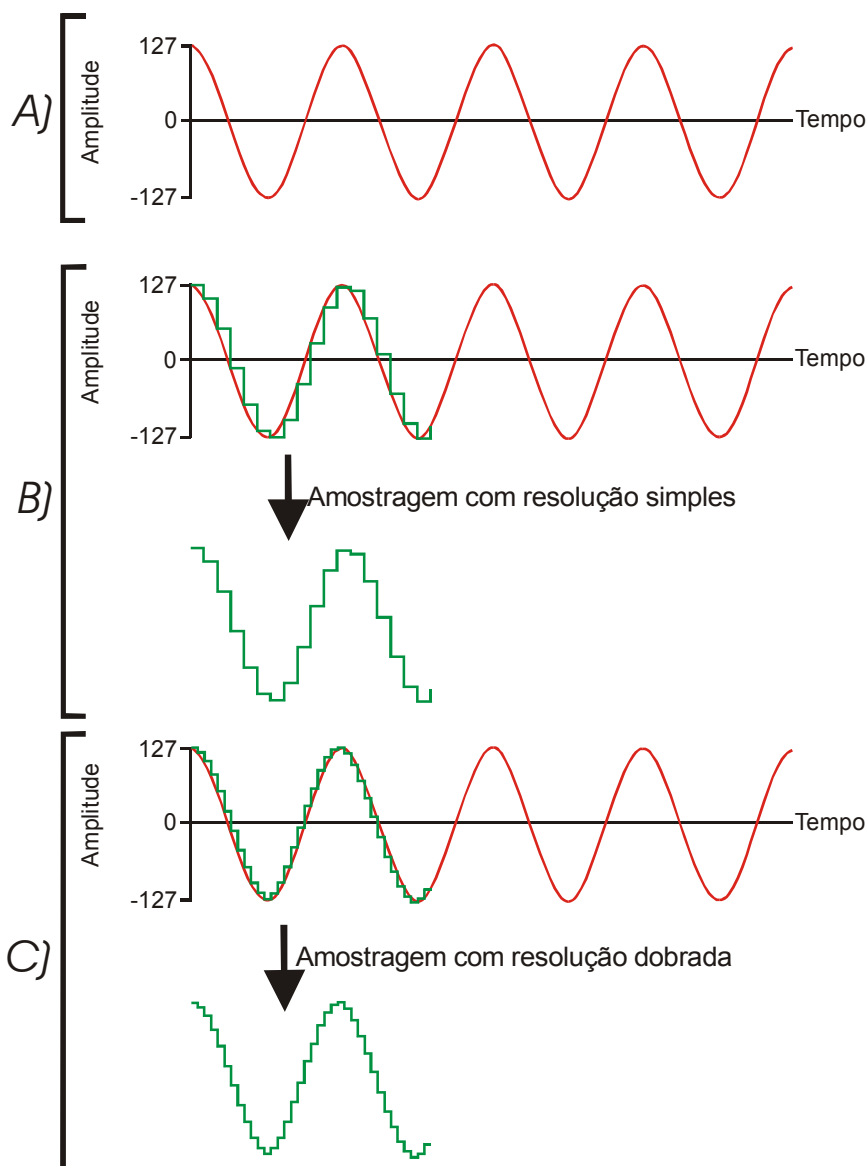


Figura 12 - O processo de amostragem digital (*codec* PCM) do som em A, B e C (neste exemplo amostras com 8 bits, resultando em aproximadamente 256 passos de quantificação: -127 até +127). Em C a taxa de amostragem é aumentada, ou seja, o intervalo de amostragem é diminuído.

Fonte: Baseado em BEGGS & THEDE (2001).

**Codificação-** conforme o valor do passo de quantificação utilizado no processo de digitalização de uma amostra, tem-se a necessidade de associá-lo a um determinado

conjunto de dígitos binários (0's e 1's). Como no caso da Figura-13 temos 256 passos de quantificação, ou seja, 256 valores possíveis, podemos codificá-los usando 8 bits ( $2^8 = 256$ ) por amostra. O tamanho das amostras vai variar conforme a necessidade de mais passos de quantificação devido a uma maior necessidade de qualidade do som ou não, e a sua conseqüente codificação com um maior ou menor número de bits por amostra (8 bits para 256 passos de quantificação, 16 bits para 65.536 passos, 24 bits para aproximadamente 16,7 milhões de passos e assim por diante). Outra medida que também é muito importante para o fluxo do som digital na Internet é a taxa de bits, que nada mais é do que o produto da taxa de amostragem pela profundidade de bits da amostra. No caso padrão da telefonia digital teríamos 8 KHz (que é a frequência amostral tida como suficiente para a voz humana pelo Teorema de **Nyquist**) de taxa de amostragem e 8 bits por amostra, o que daria  $8.000 \text{ Hz} \times 8 \text{ bits} = 64 \text{ Kbps}$ , o que é um pouco mais alto do que a capacidade de largura de banda de um modem comum de 56 Kbps (WILLRICH, 2000).

Os principais fatores que afetam a qualidade de um áudio digitalmente codificado são: o tipo de *encoder*, a taxa de bits, complexidade da música e a sensibilidade do ouvinte. Quando se é feita referência aos *codecs* MPEG (com fundamentos na Psicoacústica: ver item 3.1.1), estes fatores se tornam ainda mais preponderantes. O que conhecemos como “qualidade do som”, principalmente quando nos referimos a música, é um tanto quanto subjetivo e relativo. Em geral, a melhor forma de se analisar a qualidade do som produzido por um *codec* é através de testes com ouvintes de grande sensibilidade (com grande espectro de captação de frequências) e bem treinados.

Toda esta estrutura de dados descrita anteriormente vai ser acondicionada na forma de um arquivo de dados de áudio. Basicamente um arquivo de áudio tem a seguinte estrutura (Figura-14):

- a) cabeçalho: usado para guardar informações do arquivo como resolução, taxa de amostragem, número de canais e tipo de compressão;
- b) área de dados: os dados de áudio propriamente ditos;



c) freqüentemente um “invólucro” é utilizado para adicionar características como gerenciamento de direitos autorais ou capacidade de *streaming*.

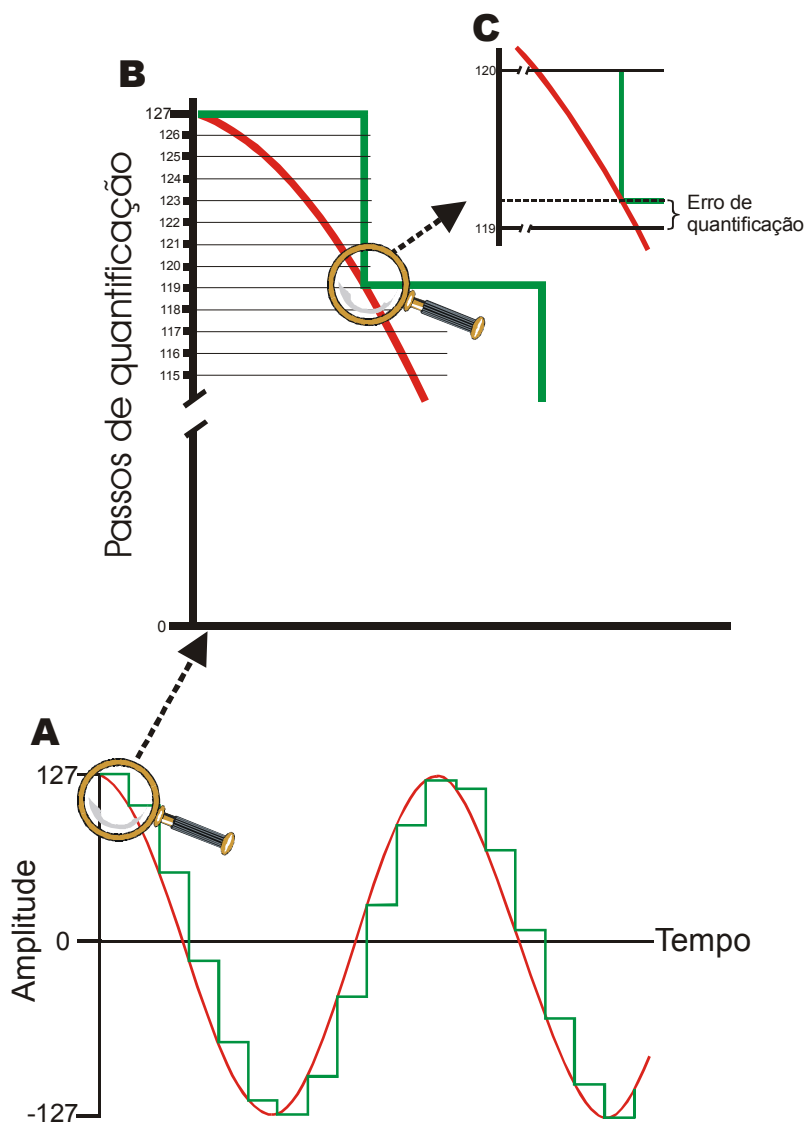


Figura 13 - Em A,B e C o processo de quantificação visto em detalhamento crescente (*codec* PCM). Observe que em C a linha da amostra (intervalo amostrado) não coincide com o passo de quantificação 119, resultando daí um erro: o erro de quantificação. A linha curva representa o sinal analógico.

Fonte: modificado de HASS (2002).

Podemos explicar a teoria de Nyquist através de um exemplo. Considere um sistema de áudio digital com uma taxa de amostragem de 44,1 KHz gravando uma onda pura de frequência crescente. Em frequências mais baixas, o som é amostrado com muitos pontos por ciclo. À medida que a frequência do som aumenta, os ciclos ficam mais curtos e cada vez menos pontos amostrais estão disponíveis para representá-los. Na frequência de 22,025 KHz, somente dois pontos amostrais estão

disponíveis por ciclo e então assim se estaria no limite do que Nyquist dizia ser possível como condição mínima ao processo. Ainda os dois pontos amostrais são adequados, pelo menos teoricamente, para recriar as frequências contidas no som original durante a conversão digital para analógica (figura-15). Contudo se a frequên-

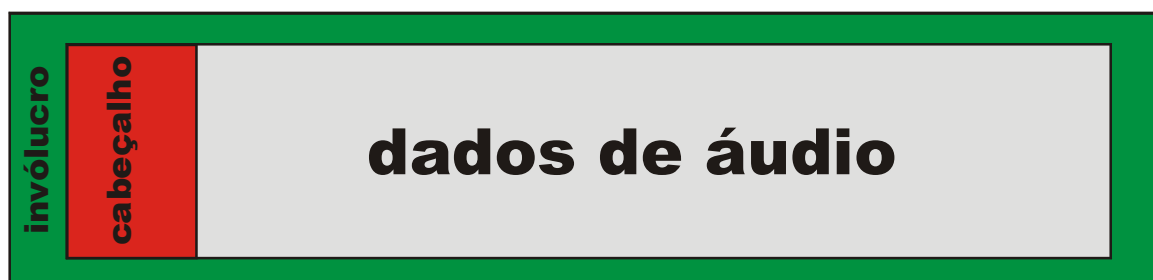


Figura 14 - Estrutura básica interna de um arquivo de som digital.

cia continua a aumentar, o número de pontos amostrais por ciclo não é mais adequado para descrever tal onda. Esta descrição inadequada pela presença insuficiente de pontos amostrais acaba gerando descrições equivalentes a frequências menores (EARLEVEL, 2002). Esta limitação é denominada de **aliasing** (Figura-16) .

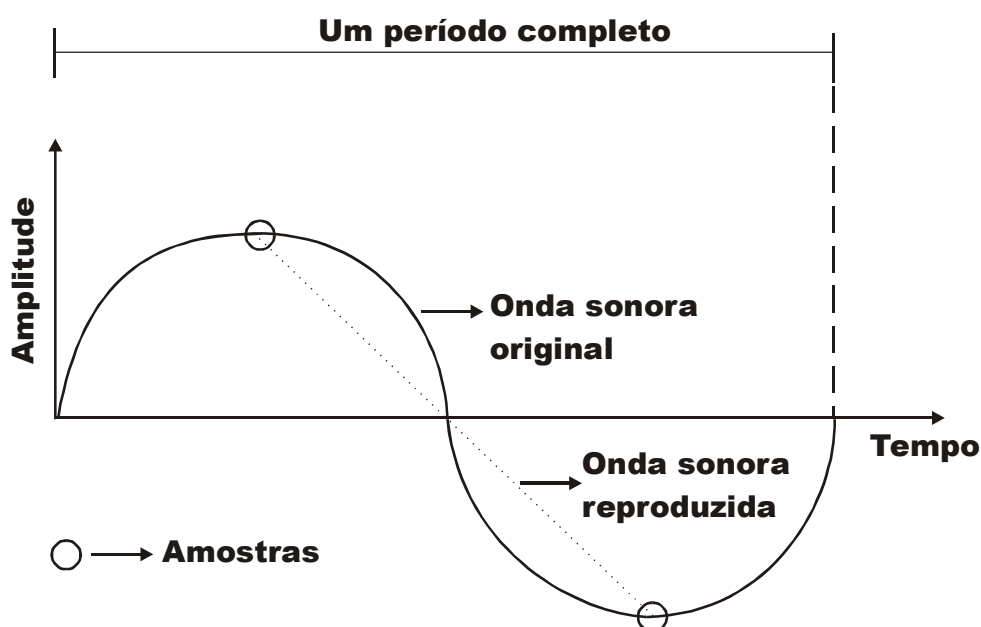


Figura 15 - O limite teórico proposto por Nyquist: 2 pontos amostrais por ciclo ou período da onda sonora. Menos do que isso provocaria o *aliasing*.

Fonte: adaptado de (<http://ei.cs.vt.edu/~netinfo/backupcs1604/sound.html>).

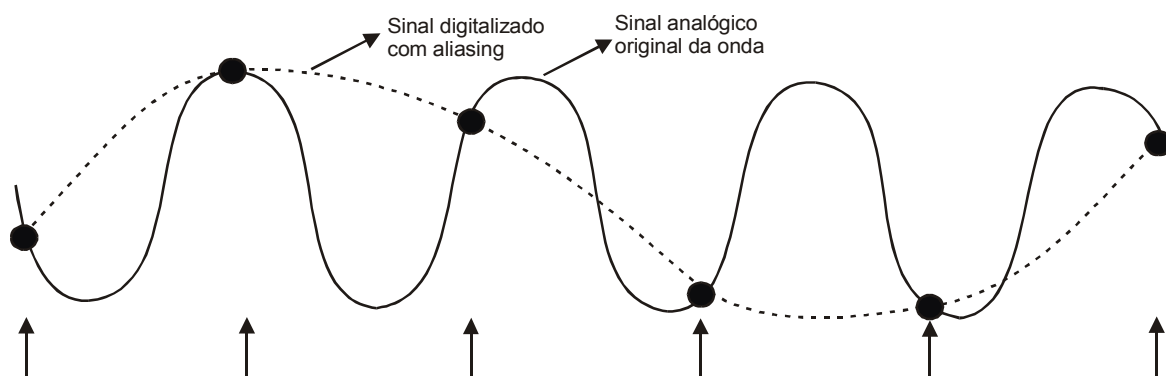


Figura 16 - Com apenas um ponto amostral (as setas indicam os pontos amostrais) por ciclo, não podemos representar corretamente as frequências presentes na onda analógica original, ocorrendo um erro de captação de frequências denominado *aliasing*. Para captar o limite de 20 KHz audíveis aos seres humanos é que padronizou-se 44,1 KHz para a amostragem dos CD's de música (pouco mais do dobro da capacidade humana).

Fonte: HASS (2002).

Poder-se-ia então resumir o processo de digitalização do som utilizando uma imagem corriqueira do mundo atual: um homem e seu computador (Figura-17). Um som produzido por um instrumento musical é captado por um microfone (A), este converte a onda sonora em sinais elétricos que são amostrados e digitalizados pelo *chip* conversor analógico-digital (B) da placa de som, sendo o som digitalizado arquivado no disco rígido. Para o homem ouvir o som digitalizado, o conversor digital-analógico deve converter os dados digitais novamente em pulsos elétricos (C), que são enviados às caixas acústicas, onde são convertidos em ondas sonoras novamente (D).

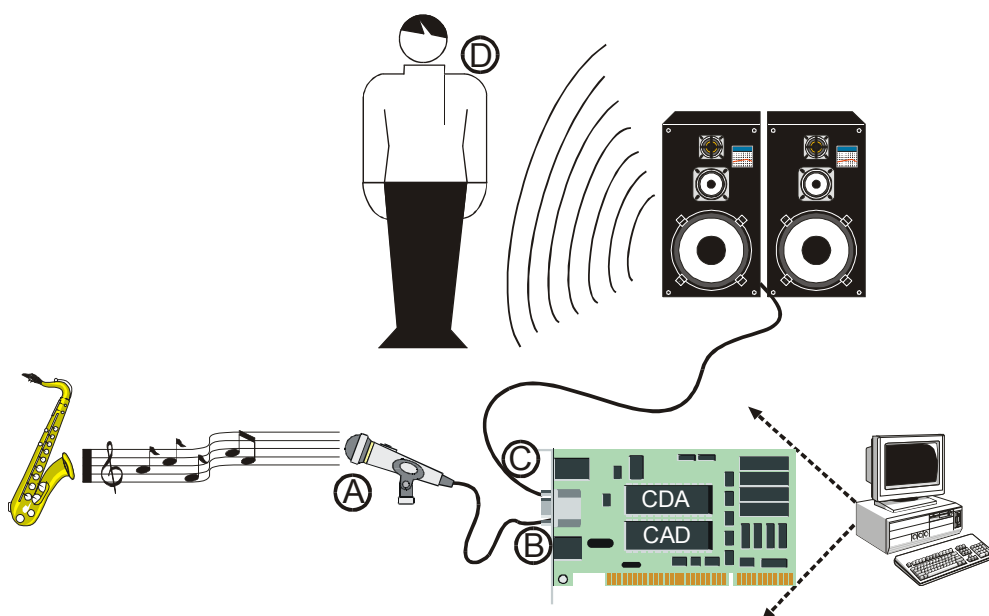


Figura 17 - O processo de digitalização do som em uma cena corriqueira do mundo real.

## 2.4- Vantagens e Desvantagens do Som Digital

O som digitalizado traz várias vantagens:

**-padrão de forma de representação do som:** o som digital, como qualquer outro tipo de informação digitalizada, não passa de um conglomerado de números, em nível mais básico, zeros (0) e uns (1), os conhecidos bits. Uma vez na forma de bits, que é a linguagem universal dos computadores, o som pode ser manipulado com o uso de qualquer sistema computacional. Desta forma o som está pronto para ser integrado com outras mídias (vídeo, imagem e texto) também digitalizadas;

**-facilidade de processamento ou edição:** estando na forma digital o som está pronto para ser manipulado e editado dos mais variados modos. O número de sistemas de *hardware* e *software* para filtragem, efeitos, mixagem do som entre outras capacidades mais, é atualmente bem grande. Facilmente são encontrados *softwares* poderosos hoje em dia acessíveis ao usuário caseiro e amador;

**-qualidade do som:** os sinais digitais são mais tolerante a ruídos e interferências que os analógicos. Como existem apenas dois níveis de sinal (alto=1 e baixo=0) tem-se que se o ruído estiver dentro de um certo limiar ainda assim o sinal será reconhecido corretamente. Além disso, o sinal analógico acumula erros a medida que é processado ou repassado de uma mídia para outra, como por exemplo, gravar um disco LP em uma fita K7. Contrariamente a isso o sinal digital é sempre reconstruído sem perda quando repassado de uma mídia digital para outra (exemplo: repassar uma música de um disco CD para uma fita DAT);

**-direitos autorais:** técnicas de criptografia, codificação e inclusão de dados sobre direitos autorais junto aos arquivos de som pode elevar bastante o nível de segurança do produto;

**-armazenamento:** o som na forma digital pode ser armazenado basicamente em qualquer dispositivo de armazenamento para computadores, como qualquer arquivo comum de texto: CD-ROM, disco-rígido, cartão de memória, fita DAT, disquetes, discos ópticos, ZIP *disks* e mais algumas dezenas de tipos de mídias;

**-transmissão:** basicamente a transmissão de som pela Internet é uma revolução. Podemos ter uma rádio, criada por uma só pessoa, em um fundo de garagem transmitindo ao vivo para o mundo todo. O Internet-fone colocou o custo da ligação internacional ao nível do custo local. Bancos de sons pelo mundo todo agora estão acessíveis a uma grande parte da população mundial e a troca de arquivos musicais virou uma mania.

Além desses fatores, tem-se também o fator durabilidade. Cada vez que um disco de vinil ou uma fita K-7 são tocados, pedaços microscópicos de vinil ou grânulos de óxido são arrancados, respectivamente.

Apesar das vantagens serem muitas, o som digital não está isento de desvantagens. Basicamente as desvantagens são:

**-a introdução de distorção** no processo de digitalização através de insuficiente taxa de amostragem, pouca profundidade de bit ou resolução da amostra e problemas inerentes ao próprio codificador (*codec*). Se a amplitude do sinal analógico que está sendo digitalizado excede a amplitude máxima que pode ser amostrada numericamente pelo conversor analógico-digital, então o sinal digital será uma versão cortada (com *clipping* ) do som em questão (Figura-18). Em relação ao problema inerente ao *codec*, pode-se citar aqui o exemplo do formato MP3, que apesar de ter um eixo de formato comum, apresenta atualmente muitas variantes do *codec* original do grupo Fraunhofer (detentor dos direitos sobre a codificação MP3). Alguns *codecs* introduzem mais distorções do que outros. Estas distorções são especialmente conhecidas pelos admiradores do som MP3. São aqueles estalos e chuviscos conhecidos pelos nomes de "*pops*" e "*glitches*" que desagradam tanto os aficionados do ramo. Obviamente, os codificadores estão evoluindo dia a dia e a qualidade cada vez tem melhorado ;

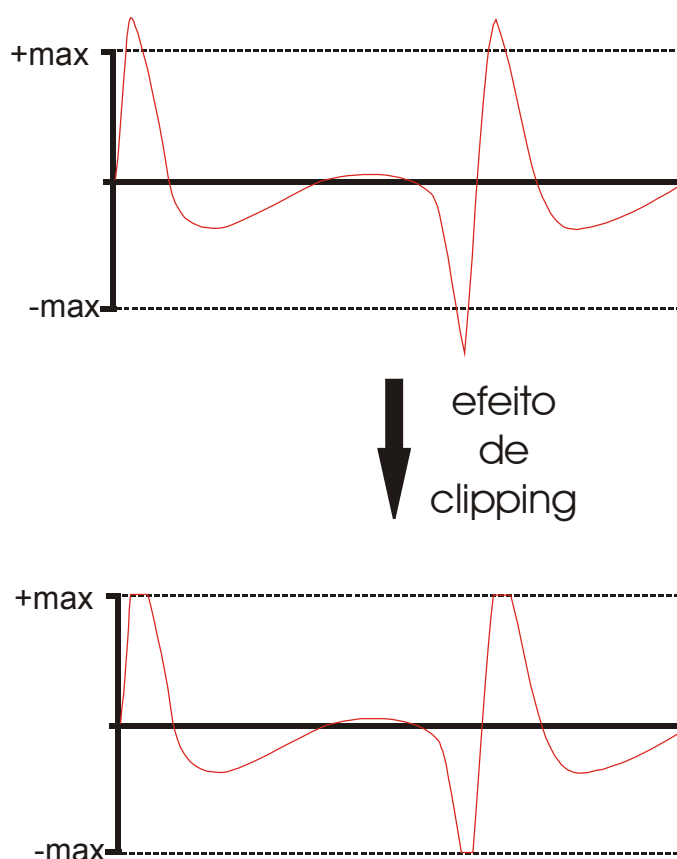


Figura 18 - Efeito de *clipping* em um sinal sonoro com amplitude alta demais para o conversor analógico-digital.

Fonte: DOBRIAN (2001)

- **o espaço necessário para armazenar** os arquivos de som pode se tornar um tormento quando estes estão em grande quantidade, sendo a necessidade de compactação quase obrigatória, caso contrário só com o uso de fitas digitais de grande capacidade para o armazenamento de tantos dados ou o uso de DVD's de várias camadas. Apenas para se ter uma idéia; 1 minuto de som digital de qualidade CD (44.1 KHz, 16 bits, estéreo) gera um arquivo de 10 Mb na forma descompactada (WILLRICH, 2000).

## 2.5- A Revolução da Tecnologia Streaming

Para se ter uma idéia da dimensão das rádios WEB, em 1998 nos EUA já havia quase 700 rádios com transmissão *on-line*. Alguns especialistas já afirmam que, ou as rádios convencionais tornam-se também *on-line* ou estarão fadadas a sumir. Um

bom *site* onde podemos ver as funcionalidades do *audio-streaming* é em **[www.audionet.com/](http://www.audionet.com/)**, onde temos transmissões esportivas, rádios-WEB, livros de áudio e um grande banco de arquivos sonoros (LIEB, 2001).

Basicamente existem dois tipos de *streaming*: *streaming* sob demanda (dados em arquivos) e *live streaming* (transmissão de fatos em tempo real).

Grande parte do *streaming* atualmente se faz usando-se o protocolo padrão da internet (TCP: *Transmission Control Protocol*), contudo ele apresenta correção de erro, sendo que durante o processo de correção o *streaming* é parado causando pequenos travamentos no processo. Outro protocolo que está sendo usado é o UDP (*User Datagram Protocol*) que não faz correção de erro, os pacotes de dados ruins são apenas descartados, sendo que no *cômputo* geral na maioria das vezes não causa alterações notáveis no fluxo do som. Contudo um problema do UDP é que ele nem sempre funciona com *firewalls*. O RTSP (*Real-Time Streaming Protocol*) é um novo protocolo desenvolvido especificamente para simplificar e melhorar o processo de *streaming* (GEISLER, 2002).

Muitos produtos como o RealAudio exigem para o gerenciamento dos fluxos de áudio (controle da quantidade, qualidade dos *streams* de áudio e de quem está requerendo o arquivo em questão) a instalação de um servidor de áudio para *streaming*. O custo destes servidores, em geral, são elevados. Outros produtos mais simples não precisam de servidores, sendo necessário apenas um *encoder* e um *player* (GEISLER, 2002).

Entre os mais conhecidos sistemas de *streaming* temos: RealNetwork's RealAudio, Apple's Quicktime, Shoutcast e Windows Media Audio (WMA). Existem outros como: Vdolive, EchoCast, Internet Wave, StreamWorks, Vosaic, VivoActive, InterVU e RadioDestiny.

Já existem atualmente sistemas para *streaming* de áudio que não necessitam de *players* para serem ouvidos nem caros servidores. Um bom exemplo é o Clipstream que usa *codec* proprietário e protocolos, *softwares* e *hardware* padrões de cliente/servidor amplamente implementados, para rodar os seus aplicativos Java. Este formato utiliza um processo intuitivo e rápido para a implementação de *streaming* em uma página WEB, *e-mail*, *banner* ou uma aplicação WEB (CLIPSTREAM, 2002).

## 2.6- Os Algoritmos de Compactação-descompactação (*CODECS*)

Em geral, nos textos de divulgação que tratam sobre o assunto, um conceito tende a ser erradamente difundido, a confusão entre o conceito de formato de arquivo e tipo de compressão usado com ele. Existem formatos, como o formato WAVE do Windows (.WAV), que suportam vários tipos de *codecs* (IMA-ADPCM, GSM, Microsoft ADPCM, PCM, A-law,  $\mu$ -law, MP3 e Divx áudio, entre outros), enquanto outros formatos suportam apenas os seus *codecs* dedicados, como o Windows Media (WMA) e Real Audio (RA). É preciso ter em mente que o núcleo de um arquivo de som é o seu método de compressão-descompressão ou codificação-decodificação (*codec*). Assim é preciso diferenciar um arquivo .WAV com *codec* MP3 de um arquivo puramente MP3 com um *codec* padrão MP3(MPEG-1 layer III).

A quantidade de *codecs* de som existentes hoje no mercado, sejam eles de uso livre ou de uso proprietário, é variada, sendo difícil precisá-la. O problema está em discernir o que é um algoritmo original e o que é uma variante.

Entre os mais conhecidos e utilizados, poderiam ser destacados os seguintes:

### **a)mu-law:**

O formato de codificação mu-law é um padrão internacional usado em telefonia, também conhecido como padrão ITU (antigamente CCITT) G.711. Ele codifica cada amostra de 16 bits em 8 bits utilizando uma tabela logarítmica. A codificação e decodificação é muito rápida e o suporte é universal. Há um *codec* variante chamado A-law usado nos sistemas de telefonia europeus (FRANKE, 2001).

### **b)PCM (Pulse Code Modulation):**

*Codec* sem áudio compactado, tornou-se o padrão em arquivos de som digital. Praticamente lido por qualquer *software* de áudio moderno. Equivale ao arquivo *raw-text* (.txt) do mundo dos processadores de texto. É usado como padrão do CD-A (CD musical comum). Também tornou-se base para codificar os formatos WAV (Windows) e AIFF (Apple) (FRIES, 2000).



**c)DPCM (Differential Pulse Code Modulation):**

É uma forma simples de compactação com perda que armazena somente a diferença entre amostras consecutivas. Usa 4 bits de resolução para armazenar a diferença não importando a resolução do arquivo original, sendo assim, um arquivo original de 8 bits terá uma compressão de 2:1 enquanto um de 16 bits terá de 4:1 (FRIES, 2000).

**d)ADPCM (Adaptive Differential Pulse Code Modulation):**

É similar ao DPCM só que o número de bits usado para armazenar a diferença entre as amostras varia dependentemente da complexidade do sinal. Existem muitas variantes deste *codec*, que é sem dúvida um dos grandes padrões mundiais. Há o padrão IMA (*Interactive Multimedia Association*), os padrões ITU G.726 e G.727, e versões proprietárias da Microsoft, Creative Labs, Yamaha e Oki. Existe ainda o ADPCM sub-banda (G.722) que é usado para áudio em linhas telefônicas ISDN (FRIES, 2000).

**e)GSM (Global System for Mobile Telecommunication):**

É um padrão internacional utilizado em telefonia celular digital. Ele usa codificação linear predictiva para comprimir substancialmente os dados predizendo a provável forma da onda sonora e registrando a diferença entre a onda predita e a forma real da onda. Existem *codecs* melhores entre os de baixas taxas de bits, contudo o algoritmo GSM está disponível gratuitamente resultando na difusão de usos em vários produtos, tais como os Internet-fones: CyberPhone, NetPhone e Speak Freely. É indicado apenas para a codificação de voz, e, apesar de críticas, está sendo atualmente bem utilizado em celulares (FRANKE, 2001)

**f)InterWave:**

É um formato de codificação proprietário criado pela empresa Vocaltec. Ele é feito especificamente para áudio em tempo real na Internet e possui grande rapidez na codificação. Ele também é a base para o produto Internet Phone da VocalTec (SERIOUS CYBERNETICS, 2003).

**g)TrueSpeech:**

É um *codec* proprietário criado pela empresa DSP, Inc. É desenhado para uso em telefonia digital (tal como o WebPhone) sendo implementado em *hardware* através de chips dedicados. O decodificador pode tocar todos os formatos TrueSpeech, mas o *software* codificador apenas gera atualmente o formato 8.5. É um dos *codecs* atualmente utilizados para o formato de arquivo WAVE da Microsoft. Os arquivos podem ser comprimidos até cerca de 30:1 dando alta qualidade e pequena largura de banda de áudio em *streaming*. Os *players* e codificadores são gratuitos e estão disponíveis para a plataforma Windows e Macintosh. Também nenhum *software* especial é necessário para gerar o arquivo TrueSpeech na Internet e não há taxas de licenciamento (SERIOUS CYBERNETICS, 2003).

**h)RealAudio:**

Este é um *codec* proprietário que é a base do formato de arquivo de mesmo nome: RealAudio. Foi o primeiro grande padrão para o *streaming* de áudio na WEB. Apesar da forte investida da empresa Microsoft e seu *codec* Windows Media Audio (WMA), o *codec* RA continua sendo muito utilizado, sendo ainda um dos padrões da grande rede. Isto é facilmente visível visto a grande quantidade de rádios-WEB transmitindo em seu formato (FRIES, 2000).

**i)ToolVox:**

*Codec* proprietário criado pela empresa VoxWare para trabalhar exclusivamente com voz e não com música e efeitos de som. Alcança grandes níveis de compressão. É a base para produtos como: TeleVox, ToolVox for Multimedia e ToolVox RT (SERIOUS CYBERNETICS, 2003).

**j)Ogg Vorbis**

*Codec* de código aberto e gratuito criado por programadores aficionados em música como um meio de evitar o controle mundial dos *codecs* de áudio por grandes corporações. Grande expectativa paira sobre este *codec* visto novas implementações poderem ser feitas por programadores do mundo todo e de várias maneiras. Contudo,

apenas o tempo poderá confirmar ou não o seu sucesso (CENTRAL DA MÚSICA, 2002b)

#### **k) ASPEC:**

É um dos algoritmos de maior qualidade de som. Pode produzir som em qualidade CD e suporta diversas taxas de bits variando de 128 Kbps pra baixo. Princípios da Psicoacústica são utilizados neste *codec*. Os *codecs* ASPEC e MUSICAM foram usados no desenvolvimento do conhecido *codec* MPEG-1 camada III (MP3) (WIKIPEDIA, 2003).

#### **l) ePAC (Enhanced Perceptual Audio Coding):**

Desenvolvido pela Lucent Technologies, este *codec* é um grande avanço em relação ao PAC original. Este rápido *codec* suporta encriptação, marca d'água e som *surround* 5.1 canais. É um formato robusto do qual se espera boa aceitação do mercado (5DDESIGN, 2001).

#### **m) ATRAC3 (“Adaptive Transform Acoustic Coding 3”):**

*Codec* proprietário da Sony. Inicialmente desenvolvido para o *Walkman* digital da Sony. Tem um fator de compactação de 12:1 e é baseado no ATRAC original. É utilizado no dispositivo MiniDisk, também desenvolvido pela Sony (5DDESIGN, 2001).

#### **n) AC-3 (Dolby Digital Multichannel Perceptual Coding ):**

Poderoso *codec* para osom *surround* 5.1 canais que é utilizado como som na HDTV (televisão de alta definição), em muitos DVD's e também nos sistemas de *Home Theater*. Desenvolvido pela empresa Dolby, pode se tornar um dos formatos mundiais para distribuição de músicas. Ele já é implementado no formato Liquid Audio (5DDESIGN, 2001).

#### **o) LPC-10E:**

LPC-10E (*Linear Prediction Coder -Enhanced*) é uma especificação americana que tem uma taxa de bits de 2400 bits/s. O *codec* adapta a voz humana a um modelo analítico simples do trato vocal, eliminando a voz e incluindo os parâmetros do

modelo mais adaptado. Depois o decodificador LPC usa estes parâmetros para gerar uma voz sintética que é mais ou menos similar a original. O resultado é compreensível, contudo acaba se assemelhando a um som robótico (BAGWELL, 2001b).

**p) CELP:**

CELP (*Code Excited Linear Prediction*), tal qual o LPC-10E, é uma especificação americana que tem uma taxa de 4800 bits/s. O *codec* CELP tem o mesmo modelo do LPC-10E mas processa os erros entre a voz original e o modelo sintético adaptado, transmitindo ambos juntos com uma representação muito comprimida dos erros (a representação comprimida é um índice compartilhado entre os codificadores e decodificadores). Este *codec* fez parte do sistema RealAudio de *streaming* em seu início, sendo que este possui agora *codec* totalmente proprietário (BAGWELL, 2001b).

**q)MACE (Macintosh Audio Compression and Expansion codec):**

Útil como *codec* de propósito geral. Foi incluso no sistema operacional do Macintosh, parte do gerenciamento de som, por muitos anos. Pelo motivo dele ter sido disponibilizado com o formato QuickTime 3.0 e posterior, ele passou também a estar acessível à plataforma Windows (BAGWELL, 2001b).

**r)Windows Media Audio:**

*Codec* proprietário da empresa Microsoft de múltiplas características e para vários objetivos como venda de músicas e transmissão de rádio pela Internet. Atualmente é o formato que mais cresce, não devido somente a características técnicas, mas também pelo poderoso *marketing* da Microsoft (5DDESIGN, 2001).

**s)MPEG (Moving Pictures Experts Group):**

Grupo de padronização mundial de *codecs* para imagem, vídeo e som. Apoiado por empresas de forte poder tecnológico e econômico. Entre os *codecs* mais famosos estão o MP3 (MPEG-1 camada III), o MPEG-2, o MPEG-2 AAC e o MPEG-4. Baseado em seus *codecs* foram desenvolvidos muitos outros *codecs* derivados proprietários. É o grupo de *codecs* de maior sucesso atualmente e muitas tecnologias

se baseiam neles: DVD, TV digital e venda de músicas pela Internet (Liquid Audio), entre outras (DUKE, 2001).

**t)VQF (Vector Quantization Format) :**

*Codec* de tecnologia diferente, patenteado pela empresa Yamaha, direcionado a taxas de bits mais baixas (abaixo de 128 Kbps) (5DDESIGN, 2001).

**u)QDX (QDesign):**

Formato proprietário desenvolvido pela empresa QDesign para áudio exclusivo para Internet, tanto para *download* completo, quanto para *streaming*. Direcionado para baixas taxas de bits (5DDESIGN, 2001).

**v)Codecs com compressão sem perda:**

Contradizendo de certo modo as tendências, os *codecs* sem perda, apesar da baixa compactação (50-60%), têm tido boa aceitação entre os aficionados da música. Este grupo é composto, em geral, de pessoas mais exigentes que se utilizam de linhas rápidas para trocar seus arquivos de áudio (cabo, ADSL, T1). Foram justamente estes dois fatores: ouvidos exigentes e linhas rápidas, junto com o barateamento e aumento da capacidade dos discos rígidos (80, 120, 200 *gigabytes*) que acabaram por firmar uma base sólida dos *codecs* sem perda no mercado. Em geral, a diferença entre eles é bem menor do que entre os *codecs* com perda, onde a diferença de tecnologias é bem mais nítida. Entre eles, podem ser citados:

- **Huffman:**

Famoso *codec* sem perda. Em qualquer arquivo de áudio existem padrões que se repetem, alguns mais que outros. Estes padrões podem ser codificados com símbolos para economizar espaço. Deste modo, no *codec* Huffman os padrões mais comuns são marcados com códigos mais curtos. Seria equivalente a trocar todas as palavras em um texto por números usando os menores números para representar as palavras mais comuns. Na realidade o *codec* Huffman não é um formato como chegam a ser Monkey's Audio e LPAC, mas sim, trata-se de um *codec* que pode ser embutido em vários formatos, entre eles, o MP3 (FRIES, 2000).

- **Monkey's Audio :**

*Codec* sem perda, *freeware* criado por Matthew T. Ashland, que é atualizado regularmente no site [www.monkeysaudio.com](http://www.monkeysaudio.com). É considerado um dos mais eficientes, atingindo maiores compactações com mais rapidez. Seus arquivos apresentam a extensão “.ape”. Este *codec*, que também acaba por definir um formato, suporta ID3 Tag (metadata) como no MP3 (RADIFIED, 2003).

- **FLAC ( Free Lossless Audio Codec):**

*Codec* de código aberto e multiplataforma, ao contrário do Monkey's Audio, que não é aberto e só existe para o Windows. Tem a extensão “.fla” e foi criado por Josh Coalson (RADIFIED, 2003).

- **LPAC ( Lossless Predictive Audio Codec):**

*Codec* de código fechado, porém gratuito, criado por Tilman Liebchen. Arquivos comprimidos com LPAC recebem uma extensão “.pac”. Bom suporte multiplataforma, porém a codificação é mais lenta (WHITTLE, 2003).

- **Shorten:**

Criado por Tony Robinson. Há uma versão gratuita básica e uma comercial com mais recursos. Realiza codificações em lote no modo pago, sendo um *codec* rápido (RADIFIED, 2003).

- **WavePack:**

Com extensão “.wv”. Muito rápido e de boa qualidade. Somente plataforma Windows. Criado por David Bryant (RADIFIED, 2003).

Ainda existem outros *codecs* sem perdas que são bem conhecidos, entre eles temos: WaveZip, WavArc (criado por Dennis Lee), Pegasus SPS (criado por jpg.com), Sonarc 2.1i (criado por Richard P. Sprague), AudioZip (criado por Lin Xiao) e RKAU (criado por Malcolm Taylor) (WHITTLE, 2003).

## 2.7- Tecnologias Variadas de Som na Internet : Internet-fone, Rádio-WEB, Videoconferência e *Voice-mail*

A videoconferência difere do áudio e vídeo *streaming* porque a comunicação é bidirecional e os atrasos (*delays*) ponta-a-ponta devem ser muito baixos (<200 ms) para a efetivação de interatividade. Os sistemas de videoconferência são bem mais amadurecidos dos que os de *streaming*, sendo oriundos do trabalho da ITU (International Telecommunication Union) na forma dos padrões H.3xx, do IETF (*Internet Engineering Task Force*) em conjunto com o MBone (*multicast backbone*) (WILLEBEEK-LEMAIR, KUMAR, SNIBLE, 2001).

O áudio em videoconferência pode usar vários tipos de codificações, entre elas, as recomendações ITU para codificação de voz (baseadas na sua maioria no PCM e no ADPCM) e outras codificações proprietárias (PISTORI, 1999).

Uma das tecnologias de áudio que despontou na Internet foi o grupo dos Internet-fones. Este tipo de tecnologia causou uma revolução na telefonia e também a antipatia das operadoras telefônicas, pois os telefonemas internacionais passaram a pagar apenas o impulso local. Para a implementação correta desta tecnologia há apenas a necessidade de uma razoável conexão com a Internet, como uma a cabo de 128 Kbps (NETPHONES, 2002).

A quantidade de dispositivos “Internet-fones” que existe já demonstra a grande aceitação no mercado que eles têm, dentre eles temos: 4eCalls, a2zconnect, BuddyPhone, Callrewards, Click2Asia, Dialpad, FireTalk, FlightComms, go2call, hotTelephone, HotVoice, iConnectHere, IP800, IPWhisper, Iscom, MediaRing, MSNMessenger, MyFreeLD, Net2Phone, PalTalk, PhoneBooth, PhoneFree, Poptel, PowWow, ReallyEasyVoice e Roger Wilco. Muitos usam *codecs* já padrões e outros usam *codecs* proprietários mais específicos (NETPHONES, 2002).

Existem duas formas de transmissão de rádio pela Internet, a transmissão digital pela WEB de rádios abertas comuns ao vivo e também as puramente WEB-rádios, das quais, muitas não passam de serviços de *streaming* de arquivos de áudio pré-programados. As rádios do segundo tipo podem ser até criadas por um usuário caseiro com a montagem de um servidor *streaming* como o Shoutcast (arquivos MP3). O primeiro tipo de rádio envolve o *streaming* em tempo real e o segundo tipo na sua maioria não. Muitos *sítes* dedicados a músicas oferecem aos usuários a

liberdade de criarem versões de uma rádio-WEB pessoal com a seleção dos estilos musicais.

A quantidade de rádios pela Internet é desconhecida. Pelo que se pode observar, já parece ultrapassar a casa de dezenas de milhares pelo mundo todo (THEODRIC TECHNOLOGIES LLC, 2002).

Outra tecnologia que também faz bom uso do áudio na Internet é o voice-mail, sistema de envio de *e-mails* sonoros. Tal qual os Internet-fones também aqui existem vários *softwares* de *voice-mail*.

## 2.8- O Fenômeno MP3, o Problema da Pirataria e Direitos Autorais

Afora a alta taxa de pirataria dos CD's musicais, que pode chegar a um em cada cinco CD's vendidos, temos também dezenas de milhões de arquivos de música disponíveis na Internet, sendo a grande maioria deles em formato MP3 (CENTRAL DA MÚSICA, 2002d).

A revolução com o uso do formato MP3 foi se estabelecendo através de vários acontecimentos. Desde o início de sua aparição no mercado, o *codec* MP3 se tornou polêmico e ligado aos problemas de direitos autorais. Em linhas gerais podem ser citados os seguintes acontecimentos que deram origem ao fenômeno MP3:

- a) grande poder de compactação com boa qualidade do *codec*;
- b) aparecimento do *software* de troca de músicas Napster, e posteriormente outros na mesma linha de uso (*Peer-to-Peer*);
- c) desenvolvimento do WinAmp, primeiro *software player* para MP3 de sucesso no mercado;
- d) lançamento pela empresa Diamond do primeiro MP3-*player* (*hardware*) portátil; várias outras marcas e tipos viriam logo após.

Em 1997 o *site* "MP3.com" foi fundado por Michael Robertson. Nos 12 meses iniciais este se tornaria o *site* musical mais visitado da Internet. Ao mesmo tempo o primeiro *player* portátil de MP3 estava sendo projetado por uma firma coreana. Assim estava lançado o MP3, o *codec* que mudaria o rumo do som digital (SMARTCOMPUTING, 2001).



Talvez o principal acontecimento para a difusão do formato MP3 tenha sido o *software* Napster. O Napster nasceu em 1999, sendo uma criação de um estudante norte-americano de 18 anos chamado Shawn Fanning. O Napster era um *software* de troca de arquivos tipo usuário-usuário (*peer-to-peer*) sem a necessidade de um *site* centralizador. O que havia, era apenas um índice central dos arquivos musicais. O sucesso do Napster foi grande, sendo que em meados de 2000 já havia mais de 40 milhões de usuários (e a cada dia o número crescia mais) pelo mundo todo trocando músicas de todos os tipos pela Internet. Então, nesta mesma época, várias ações judiciais são impetradas contra a empresa Napster com a acusação de pirataria generalizada de músicas. Segue-se uma batalha judicial que levaria meses e que terminaria com a compra da Napster pelo mesmo conglomerado que é dono da gravadora BMG (PIERCE, 2003).

Entretanto, após o fechamento da Napster surgiram vários outros programas de troca de arquivo *peer-to-peer*. Basicamente estes programas clientes (KaZaA, Grokster, Morpheus, LimeWire, BearShare, ToadNode, WinMX e Gnucleus, por exemplo) usavam 3 grandes redes: Gnutella, FastTrack e OpenNap. Existem ainda redes menores. A tendência atual é de que estes sistemas sofram grandes pressões para implementação de sistemas contra pirataria. Alguns já fecharam e não atuam mais na rede. Contudo os usuários tendem a responder criando outro sistema de acesso livre. Por mais que a pressão da indústria fonográfica aumente, mesmo assim, ainda existirão as opções de trocas de arquivos por *newsgroup* e IRC (ABOUT.COM, 2002).

A atividade de pirataria musical em massa não é uma coisa nova, vindo desde a década de 1960 quando houve um grande aumento do uso de fitas-cassetes para a gravação de músicas de discos de uma forma barata (RUBINSTEIN, 2001).

Mas o que se entende por pirataria pode ser bem relativo. A RIAA (*Recording Industry Association of America*), Associação Americana das Indústrias Fonográficas, afirma que se uma pessoa copia um CD de músicas dela para uso particular, não-comercial, já está fazendo pirataria. Contudo, outra linha jurídica propõe que tendo adquirido o direito de escutar as músicas (comprado o CD original ou qualquer outra mídia que seja), o consumidor poderá fazê-lo do modo que quiser: gravar em fita k-7, copiar o CD, transpor para MP3 para acessar em seu MP3 *player* ou gravar para qualquer tipo de mídia com qualquer tipo de formato (sendo respeitada a parte legal

correlacionada aos formatos). A pirataria só se caracterizaria quando este consumidor desse a cópia a outra pessoa ou a usasse com qualquer fim comercial. Devido a estes acontecimentos as indústrias fonográficas e de tecnologia de áudio criaram o “*Secure Digital Music Initiative*” (SDMI), que é um grupo destinado a criar tecnologias e padrões para proteção de direitos autorais e combate à pirataria (FRIES, 2000).

Órgãos de pesquisa nos EUA concluíram que pessoas que têm acesso a arquivos MP3 comprem mais CD's que as pessoas que não tem tal acesso. Segundo as pesquisas, o fato das pessoas estarem tendo a possibilidade de ouvir música *on-line* e "experimentar" antes de comprar, as deixa mais propensas a adquirir CD's (CENTRAL DA MÚSICA, 2002e).

Apesar da pressão da indústria fonográfica, os *sítes* de distribuição de músicas gratuitas *on-line* continuam crescendo. O regime ditatorial das gravadoras, que em geral acabam levando grande parte do lucro devido aos artistas, está criando uma tendência mundial: a produção independente. Já existem nos EUA, processos de consumidores contra os altos preços dos CD's musicais. As gravadoras de uma forma ou de outra terão que se adaptar também aos sistemas de venda e difusão de música pela Internet, caso contrário poderão colapsar (CENTRAL DA MÚSICA, 2002g).

Depois da onda de trocas de arquivos musicais MP3, começaram a aparecer trocas de outros formatos como RealAudio, Ogg Vorbis e principalmente WMA. Na realidade os direitos autorais já eram violados bem antes, através da troca de arquivos MIDI.

## 2.9- A Disputa pelo Formato Padrão

Segundo pesquisas feitas no ano de 2000, tinha-se nesta época uma predominância do formato RealMedia (da empresa RealNetworks) para a tecnologia *streaming* de mídias na Internet. Contudo, o crescimento em uso do formato WMA da empresa Microsoft tem sido muitas vezes maior que o do formato RealMedia nos últimos 3 anos. Apesar de existirem outras empresas que atuam com a tecnologia *streaming*, a principal disputa comercial trava-se entre as empresas Microsoft e RealNetworks (SMARTCOMPUTING, 2001).

Por outro lado, o formato Ogg Vorbis tem no bojo de seu desenvolvimento a função de evitar o controle predatório de um formato proprietário na Internet, sendo que aqui podemos fazer referência ao Windows Media e ao RealAudio. O Ogg Vorbis tem assim o código aberto e livre de patentes, podendo ser adaptado e usado por terceiros (LITEXMEDIA, 2003).

Uma das estratégias da Microsoft que ajudou a firmar o formato Windows Media Audio no mercado foi o fato de que seu *player*, o Windows Media *player*, vir integrado com os sistemas operacionais Windows (me, 2000, XP). Estratégia similar foi usada antes com o Microsoft Internet Explorer, fato que deu origem a uma ação judicial movida pela empresa Netscape, então forte concorrente com o seu *browser* Netscape Navigator, na época.

## 2.10- Tipos de Conexão e Vias de Transmissão

O fator principal que guiou o desenvolvimento dos formatos de som para a Internet foi a capacidade de tráfego de dados por unidade de tempo. É o que se chama de largura de banda. Como aumentar a largura de banda tem quase sempre um custo alto: fibra ótica, novos canais de satélites e várias outras implementações em *hardware*, optou-se primariamente pela compactação dos dados para a transferência pela rede, fazendo assim com que o grande público tivesse acesso a tecnologias de áudio que são revolucionárias, como a telefonia por IP, a audioconferência, o *voice-mail* e a rede de troca de arquivos musicais descentralizada P2P (*Peer-to-Peer*).

Além da compactação dos dados foram desenvolvidos vários tipos de conexões com a Internet. Entre os principais tipos temos (FLORIDA CENTER FOR INSTRUCTIONAL TECHNOLOGY, 2002):

- 1) Modem 56Kbps (conexão análoga) – com acesso discado, já está deixando de ser o padrão nos países mais ricos. É o tipo de conexão mais lento que há (56, 33,6 e 28.8Kbps), sendo apenas indicado para troca de arquivos texto e pequenas imagens. A taxa de 56 Kbps é raramente alcançada, na realidade, a taxa fica ao redor de 40Kbps. Para o uso com animação e áudio praticamente se torna de pouca utilidade. As únicas vantagens deste tipo de conexão estão no fato dela ainda compor a grande massa de usuários da grande rede, estar a disposição em qualquer lugar com linha telefônica comum e ser suportada por qualquer provedor de WEB.
- 2) ISDN (*Integrated Services Digital Network*) - tem conexão digital oferecendo probabilidade mínima de erro em relação à conexão análoga comum. Uma linha simples com dois canais chega a 128 Kbps. Pode utilizar o cabeamento telefônico comum tendo a vantagem de ser fluxo digital de informação. A disponibilidade de acesso a esta conexão varia conforme a área em questão e sua distância da central telefônica. É necessário a instalação de um adaptador ISDN para a efetivação da conexão com o computador do usuário final. Como o modem normal, também tem o acesso discado, contudo o processo leva de 1 a 2 segundos apenas.
- 3) DSL – existem vários tipos de conexão DSL, cada uma delas com capacidades diferentes. A ADSL (*Asymmetrical Digital Subscriber Line*) tem capacidade de *download* de até 8 Mbps e de *upload* de até 1 Mbps. Outro tipo comum é o SDSL (*Symmetrical Digital Subscriber Line*) com *download* e *upload* de 1,54 Mbps. O primeiro tipo de conexão é o mais indicado ao usuário comum, justamente pelo motivo de em geral baixar-se mais arquivos do que se carregar no dia-a-dia. O sistema DSL usa o cabeamento telefônico normal para uma conexão digital rápida e constante (24 horas). O padrão ADSL exige o uso de modem específico.

- 4) Conexão a cabo – podendo chegar a 10 Mbps. Contudo para o usuário final conforme a configuração dos provedores a cabo, a velocidade varia na sua maioria de 128 até 512 Kbps. Se faz necessária a instalação de um modem a cabo. Apesar das vantagens de ser uma conexão bem rápida, também apresenta as suas dificuldades. Está limitada às áreas urbanas (principalmente as residenciais), sendo que os provedores são unicamente as operadoras de TV-Cabo. Trata-se de uma conexão constante (24 horas). A velocidade de transferência pode ser reduzida conforme a quantidade de usuários de linha a cabo pela vizinhança.
- 5) T1 e T3 – a linha T1 permite a transmissão digital de dados em até 1,54 Mbps. A linha T3 permite velocidades de até 44,7 Mbps. Esta última linha tem custo elevado e está reservada a grandes empresas, órgãos do governo, universidades e conexões entre grandes computadores.

A grande tendência é da melhoria das conexões utilizando-se das linhas telefônicas digitais (ISDN e ADSL) e também pelo sistema de TV a cabo já existente nas residências. De outro modo o custo seria muito elevado pela mudança dos cabos já existentes.

Além dos sistemas cabeados existem também as conexões via rádio e via satélite, a primeira já é um padrão, porém a segunda ainda está em fase inicial de expansão no Brasil e de uso específico a certos usuários.

## **2.11- A Importância do *Design* de Som em um *Website***

As pessoas somente retêm 20% do que elas vêem e 30% do que elas escutam. Entretanto, elas se lembram de 50% do que elas vêem e ouvem, e tanto quanto 80% do que elas vêem, ouvem e fazem simultaneamente (GERTH, 2001).

Podem ser citados basicamente três tipos diferentes de sons: fala, música e efeitos sonoros.

A importância do som em um *website* é indiscutível, os benefícios principais podem ser enumerados:

- 1) ensinar maior interesse e impacto, mantendo o público desperto, trabalhando emoções (exemplo: a alegria de um som de torcida ou a tristeza de uma marcha fúnebre);
- 2) melhor retenção do produto que se quer vender ou divulgar, esta é a filosofia dos *jingles*;
- 3) melhoria da imagem: você passa características da sua imagem (*marketing* pessoal), para seu público através de uma música (ser relaxado, refinado, clássico);
- 4) alívio cômico (exemplo: grito de Tarzã no meio de uma página muito séria sobre investimentos);
- 5) valor agregado: pesquisadores do MIT Media Lab testaram pessoas assistindo imagens idênticas em aparelhos de TV idênticos. As pessoas relataram como sendo de melhor imagem aquele que tinha o melhor som (isto feito inconscientemente).

Existem muitos fatores a ponderar quando se pretende colocar sons em uma página WEB:

- que tipo de som usar: música, voz, sons de efeito ?;
- os sons facilitam ou ajudam a compreensão da informação fornecida na página?;
- que tipo de conexão o público-alvo utiliza ?;
- há necessidade de baixar um *plug-in* para escutar o som ?;
- a qualidade do som é importante ?;
- qual a plataforma de sistema do público-alvo ?;
- o problema dos direitos autorais;

Em geral, têm-se alguns princípios básicos em se seguir ao se trabalhar com arquivos de som em um *website*, entre os principais poderiam ser citados:

- a) deve ser evitado o uso de som de fundo (*background*) na página, e quando colocá-lo for inevitável, deve ser adicionado um modo de desabilitá-lo por completo (colocar o valor de desligado em uma variável para não ficar ativando o som de fundo quando na mudança de página dentro do mesmo *site*);
- b) o tamanho do arquivo de som em um arquivo para *download* nunca deve ser esquecido (exemplo: "Hino do Flamengo-859 Kb"), caso contrário o usuário tentará baixar o arquivo e não entenderá a demora;
- c) muitas vezes, um *site* linca com um som em outro *site*; a menos que o criador do *link* tenha permissão do proprietário do *site* portador de tal som, estará se praticando roubo de largura de banda daquele *site* (pensando com um pouco de ética o autor de páginas WEB sempre deve colocar os sons dentro de seu próprio *site*);
- d) deve-se sempre tomar cuidado com os direitos autorais, até mesmo a "risada do Pato Donald" poderá ser limitada em uso por tais direitos;
- e) ao colocar um formato não-padrão deve-se adicionar *links* para se obter o correlato *plug-in* para ouvir tal som; e também deve-se adicionar uma forma alternativa para o usuário obter pelo menos a parte principal da informação que o som contém, um texto por exemplo;
- f) deve-se errar antes por omissão do que por excesso, evitar o uso exagerado de sons é sempre de bom grado.

## **CAPÍTULO 3- OS FORMATOS DE ARQUIVOS DE SOM ENCONTRADOS NA INTERNET**

Com relação aos resultados da pesquisa bibliográfica geral e análise dos formatos digitais de áudio na WEB deve-se primeiramente fazer algumas colocações. Observou-se, por ser tratar de uma tecnologia bem atual (uma grande onda de novas tecnologias de áudio ocorreu a partir de 1998) e que evolui com grande rapidez, que o material disponível para consulta baseava-se quase que totalmente na própria Internet. O número de livros encontrados sobre o tema em pesquisa foi limitado, e quando estes eram encontrados, as informações eram repetitivas ou muitas vezes superficiais e até mesmo obsoletas. Assim, maior esforço de pesquisa foi direcionado para a Internet, utilizando-se vários mecanismos de busca para a pesquisa de informações sobre o tema em questão. Primou-se pelo uso de mecanismos de busca de eficiência conhecida: Google, Altavista, Yahoo, Lycos, Excite e AOL. De nada adiantava aumentar o número de buscadores ou utilizar metabuscas, pois o resultado mostrou-se ineficiente e com muitos itens de informações sem correlação (popularmente conhecido como "lixo"). O buscador Google ([www.google.com](http://www.google.com)) se mostrou como o mais eficiente na busca. Deste modo, uma intensa pesquisa foi feita sobre os formatos de arquivos de som na Internet que são descritos aqui neste capítulo.

### **3.1- Os Principais Formatos de Arquivos de Som na Internet**

Foi apenas com o surgimento da Internet que uma grande variedade de formatos de arquivos de áudio digital se desenvolveu. Antes da Internet, os formatos eram específicos às suas respectivas plataformas de *hardware* e sistemas operacionais. Com o tempo apareceram vários arquivos multiplataformas e também programas auxiliares que capacitavam a um sistema operacional executar arquivos criados em outro. Os *softwares* editores de áudio também seguiram a tendência estabelecendo suporte a uma grande gama de formatos de áudio digital (DUKE UNIVERSITY, 2001).

Tem-se basicamente 3 grupos de arquivos de som na Internet. O primeiro deles, na realidade, não contém som real, é o que chamamos de padrão MIDI e suas variantes. Estes arquivos nada mais são do que comandos enviados a um



sintetizador (seja de um instrumento musical ou de uma placa de som de um computador) para gerar respectivas notas em determinados momentos, compondo assim uma música ou o efeito sonoro desejado no final. A renderização da música (síntese artificial) pode se dar da forma FM ou *wave-table*. Na forma *wave-table* os comandos contidos no arquivo MIDI ativam notas que são amostras (*samples*) reais de sons que estão no próprio chip do sintetizador, assim, o som final sai de uma maneira muito mais natural, sem aquele som tipo "caixinha de música" ou dos videogames antigos de 8 bits (temos aqui o tipo de síntese que ocorre na forma FM, que é uma forma totalmente artificial de síntese de som).

O segundo grupo trata-se de um grupo híbrido que tem no interior do arquivo comandos que controlam *samples* de sons (sons reais) que também estão inclusos no arquivo. Existe uma infinidade de variantes de formatos e subformatos. Como bom exemplo pode-se tomar o formato MOD, onde o criador do arquivo define uma tabela de *samples* que contenham os mais variados sons, criados por ele ou não, e depois compõe com estes *samples* através de comandos, uma música ou um som qualquer final bem mais complexo. Este grupo é na realidade uma mistura de um padrão MIDI e som real. Esta filosofia de criar sons gerou uma grande legião de adeptos no mundo todo que agora tinham o poder de mesmo não sendo músicos e sequer saberem tocar qualquer instrumento, conseguiam finalmente compor música. Arquivos híbridos como o MOD já estão no mercado há bastante tempo, enquanto outros são mais recentes e com objetivos mais comerciais e definidos, como o formato Beatnik, feito para a Internet.

O terceiro e último grupo é dos arquivos puros de sons reais (*samples*) como os conhecidos Microsoft-IBM WAVE e Apple AIFF. Este terceiro grupo também evoluiu bastante com o desenvolvimento de avançados *codecs* e da melhora da tecnologia *streaming*, fazendo com que arquivos antes imensos pudessem agora ser ouvidos e transitar por toda a rede mundial (ANDERSSON, 2001).

### **3.1.1-O Grupo MPEG**

O grupo MPEG (*Moving Pictures Experts Group*) é um comitê de padrões técnicos estabelecido em 1988 e que trabalha sob a supervisão da ISO (*Internacional*

*Standards Organization*). A proposta básica do MPEG é o de propor padrões de codificação para áudio, vídeo e imagens gráficas. Organizações do mundo todo estão envolvidas no desenvolvimento de padrões MPEG. Cada uma destas organizações detém direitos de patentes sobre qual formato ajudou a desenvolver.

Os *codecs* MPEG se baseiam nos princípios da Psicoacústica. No caso do MP3, para exemplificar, tem-se principalmente os seguintes processos e princípios:

- a) Divisão em sub-bandas de frequências do sinal de áudio para otimização dos processos a serem aplicados.
- b) A audição humana é mais sensível a faixa de frequências entre 2 e 4 KHz (faixa da voz humana).
- c) Mascaramento de sons mais suaves por sons mais altos que estão próximos dos primeiros, tanto em frequência, como também no tempo.
- d) Reserva de bits, onde os bits de certas passagens mais simples de um sinal de áudio são deslocados para passagens mais complexas, dando maior resolução no som digitalizado final onde ela é mais necessária.
- e) Diversos modos de tratar com o modo estéreo, visto que muito da informação é igual em ambos os canais. Cada modo varia no montante de compressão e na fidelidade do efeito estéreo.
- f) Codificação Huffman: aplicação final, depois dos processos psicoacústicos, de uma codificação tradicional sem perda.

Estes processos anteriores são gerais e não se aplicam necessariamente a todos os *codecs* MPEG. Contudo é regra geral que princípios da Psicoacústica sejam aplicados nas codificações MPEG (retirada de sons que a audição humana não percebe), seguidos de uma técnica de compressão de dados sem perda como a de Huffman.

Foram desenvolvidas 3 famílias de padrões: MPEG-1, MPEG-2 e MPEG-4. O MPEG-3 foi descontinuado e fundido ao MPEG-2. Ainda existem o MPEG-7 e o MPEG-21, que são padrões informacionais e estruturais (em desenvolvimento) para a multimídia, não sendo *codecs* de áudio ou vídeo. Determinando-se as principais características de cada grupo assim teremos:

**MPEG-1:** codificação em mono e estéreo, com frequências de amostragem de 32; 44,1 e 48 KHz e taxas de bits entre 32 e 448 Kbps. Apresenta 3 camadas com níveis de complexidade crescentes. Padrão aprovado em 1992.

**MPEG-2:** compatível descendentemente com MPEG-1 (ou seja, um *decoder* MPEG-2 camada I pode tocar um arquivo MPEG-1 também de camada I, e assim desta mesma forma com as outras duas camadas: II e III). O MPEG-2 possui até 5 canais normais e um canal melhorado de baixa frequência. Adiciona suporte a amostragens de 16; 22,05 e 24 KHz para taxas de bits entre 32 até 256 Kbps para a camada-I, até 384 Kbps para a camada-II e de 8 até 320 Kbps para camada-III. Padrão aprovado em 1994.

**MPEG-4:** um *codec* multi-uso desenhado para ser um padrão geral de sistemas de *streaming* multimídia (áudio, vídeo, texto, imagens, 2-D, 3-D, realidade virtual) no futuro, tanto com pouca ou com grande largura de banda na rede (de 5 Kbps a 20 Mbps). Com suporte para sistema de gerenciamento de direitos autorais. Apresenta grande poder de minimização de erros de transmissão. Permite a composição de áudio natural ou sintético. Apresenta suporte a interatividade permitindo a manipulação dos dados de áudio e dados visuais. É baseado em avanços de *codecs* e tecnologias anteriores como o QuickTime. Padrão aprovado em 1998 e 1999. Contudo na prática, parece não estar sob consenso ainda hoje em dia.

Existe um formato não-ISO chamado de MPEG-2.5, criado pelo Instituto Fraunhofer para a melhoria do MPEG áudio em baixas taxas de bits, que se fundamenta na diminuição das taxas de amostragens (8; 11,025 e 12 KHz) e manutenção de satisfatória resolução da amostra.

Existem esquemas de codificação correlacionados entre si dentro dos padrões MPEG, eles são denominados de camadas, que existem tanto no MPEG-1 quanto no MPEG-2. São 3 camadas para os dois padrões, sendo que cada camada usa a mesma estrutura básica e inclui as características da(s) camada(s) inferior(es) a ela. As camadas têm variados níveis de compatibilidade entre si. Dentro do MPEG-1, as 3 camadas são compatíveis descendentemente, ou seja, um *decoder* de camada III deve também ser capaz de decodificar um fluxo de camada I e II. As camadas MPEG-1 e suas correspondentes no MPEG-2 são compatíveis até um certo limite. *Decoders* MPEG-2 devem ser capazes de decodificar arquivos de MPEG-1 de camada correlata e os *decoders* MPEG-1 devem ser capazes de tocar os canais direito e esquerdo de um sinal MPEG-2 de camada também correlata. Camadas mais altas oferecem som progressivamente melhor em taxas de bits comparáveis e requerem *software* de codificação com maiores níveis de complexidade, exigindo maior poder de processamento para codificação e decodificação do áudio (FRIES,2000).

O *codec* AAC (MPEG-2 *Advanced Audio Coding*) foi considerado padrão MPEG em 1997 e lançado ao público em 1999. O Instituto Fraunhofer cooperou com as companhias AT&T, Lucent Technologies, Sony e principalmente a Dolby para desenvolvê-lo. Este *codec* é capaz de produzir melhor qualidade de som em menores taxas de bits do que o MP3. Ele não é compatível com as 3 camadas (I, II e III) do MPEG-2, tratando-se realmente de um formato revolucionário. Dada a sua solidez, acabou dando origem a vários outros desenvolvimentos proprietários: A2B, Liquid Audio, ISO AAC e Astrid/Quartex ACC (5DDESIGN, 2001).

O AAC representa o atual estado-da-arte em codificação de áudio. Ele é capaz de portar até 48 canais de áudio, 15 canais melhorados de baixa frequência, 15 *streams* de dados embutidos e capacidade multilínguas. Testes do grupo MPEG demonstraram que ele, em 96 Kbps, oferece qualidade levemente melhor do que o MP3 a 128 kbps, oferecendo um nível de compressão de 16:1 em qualidade próxima ao CD. O AAC trabalha em taxas de bits de 8 Kbps (voz em mono) até áudio de alta qualidade (320 Kbps). A grande flexibilidade abriu um amplo campo de aplicações para o AAC, tais como: *broadcasting*, rádio digital, DVD, comunicação em tempo real, *streaming* de áudio, vídeo e TV digital. Com taxas de amostragens entre 8 kHz e 96 kHz, o método está bem preparado para futuros melhoramentos. A força

direcionadora para o desenvolvimento do AAC foi a necessidade de uma eficiente codificação para sinais *surround* tais como os utilizados nos modernos cinemas de hoje (NUNTIUS,2002) (Quadro-2).

A empresa Coding Technologies desenvolveu um *codec* denominado AAC *Plus*, que é a combinação do MPEG AAC e de sua tecnologia SBR (*Spectral Band Replication*). SBR é uma técnica única de extensão de faixa de frequências que capacita aos *codecs* de áudio produzir a mesma qualidade de som em metade da taxa de bits. Como resultado o AAC *Plus* gera som qualidade CD (equivalente 128 Kbps MP3) com 48 Kbps. Este nível de eficiência capacita novas aplicações no mercado de difusão digital. Este tipo de tecnologia pode também, como outras mais, se tornar parte do padrão MPEG-4 (CODING TECHNOLOGIES, 2002).

O formato MP3PRO é um avanço desenvolvido também pela empresa Coding Technologies em cima do MP3 (auxiliado pelos grupos Thomson Multimedia e Fraunhofer Institute). Em testes preliminares viu-se que 95% dos ouvintes não conseguiram identificar qualquer diferença entre a qualidade de um arquivo MP3 de 128 kbps e um MP3PRO de 64 kbps, sendo que o arquivo codificado em MP3PRO tem a metade do tamanho do codificado no formato antigo. O formato MP3PRO é totalmente compatível com a base instalada para o formato MP3. Assim, os *softwares* e MP3 *players* portáteis que têm capacidade para reproduzir arquivos MP3, também poderão reproduzir o MP3PRO sem a necessidade de qualquer atualização nos mesmos, e vice-versa. Contudo existe um problema, os *decoders* MP3 normais não conseguirão tocar o arquivo no modo MP3PRO e sim no modo padrão, ou seja, o que se ouve então é um pobre arquivo de MP3 normal de 64 kbps e não um MP3PRO de 64 kbps (CENTRAL DA MÚSICA, 2002f).

### **3.1.2- MPEG- 1 camada 3 ("MP3")**

#### **3.1.2.1- Histórico**

O formato MPEG-1 *layer-III* foi desenvolvido pelo instituto de tecnologia alemão Fraunhofer, a Universidade de Erlangen e a empresa Thomson Multimedia no início da década de 1990 obedecendo a padronizações do ISO-MPEG (FRAUNHOFER

INSTITUT, 2001). Apesar de muitos pensarem que o formato é de uso aberto e livre, o MP3 precisa ser licenciado quando utilizado por outra empresa ou grupo de pesquisa. A quantidade de *encoders* baseados legal ou ilegalmente no MP3 atualmente é até desconhecida visto a grande quantidade de variantes que surgiu. O MP3, constantemente é confundido com MPEG-3 por muitos, sendo que este último é um formato completamente diferente que teve seu desenvolvimento descontinuado. O MP3 é na verdade a terceira camada (que é basicamente o nível de complexidade da codificação) de um grupo de 3 camadas do padrão MPEG-1. Esta terceira camada (MP3) é a camada com melhor qualidade de som em menores taxas de bits (THOMASON, 2001).

O sucesso do MP3 na Internet foi tão grande no início que o termo “MP3” chegou a ser o segundo termo mais procurado nos serviços de busca da WEB, atrás só da palavra “sexo” (INFO EXAME, 1999). Analistas da área de música digital afirmam que cerca de  $\frac{3}{4}$  das músicas descarregadas da WEB estão no formato MP3. O mercado de música *on-line* tende a superar o valor de 1,5 bilhão de dólares nos próximos 3 anos, porém, a resistência das gravadoras em colocar os seus produtos *on-line* é grande, pois o MP3 não tem sistema de proteção contra cópia. Do outro lado está o consumidor, que na sua maioria é contra sistemas de proteção de cópia (THOMASON, 2001).

Em resposta a Microsoft, que reclamava para si uma maior eficiência em baixas taxas de bits do seu formato para música, Windows Media Audio (.WMA), as organizações por detrás do formato MP3 lançaram em meados de 2001 o formato MP3PRO. Basicamente, este formato (cujos criadores afirmam ter a metade do tamanho do MP3 normal e a mesma qualidade) utiliza-se de uma técnica chamada de *replicação de banda espectral* que recupera as altas frequências que são descartadas no processo de codificação do MP3 normal em baixas taxas de bits (80 Kbps para baixo). Apesar deste novo formato ser compatível com os *players* antigos de MP3, a faixa de altas frequências recuperadas só pode ser ouvida em *player* próprio. Em geral, o que se espera é a aceitação pelo mercado com o lançamento de um banco musical grande neste formato. O MP3PRO possui sistema de proteção de direitos autorais (marca d'água) (LITEXMEDIA, 2003).

### 3.1.2.2- Características

Uma música de CD de áudio musical comum é gravada a uma taxa de bit de 1.411 Kbps, enquanto que uma música em MP3 de alta qualidade é criada e usada entre taxas de 128, 192 e mesmo 320 Kbps. A taxa de 128 Kbps é considerada como a mínima para "qualidade CD". Entretanto, o padrão de qualidade mínima não existe na realidade, pois depende da sensibilidade auditiva e musical de cada pessoa. Um arquivo de 128 Kbps pode parecer bom para uma pessoa e ruim para outra. Além do mais, por causa das características do *codec* MPEG, que se utiliza de ditames da Psicoacústica, o tipo de música (clássica, rock, samba) com seus instrumentos diferentes (diferentes misturas de frequências e timbres), também será determinante da taxa de bit mais adequada (DUNN, 2001).

O som MPEG usa um método de compressão que determina quais bandas de frequências aparecem mais vezes de amostra para amostra (durante o processo de amostragem na digitalização) e concentra mais a compressão nestas bandas ou faixas de frequências, utilizando ainda princípios da Psicoacústica para filtrar ruídos que não são percebidos pela audição humana (CASS & SEYMOUR, 2001).

No decorrer do uso dos arquivos MP3, e visto essencialmente a sua utilização ter sido direcionada para quase que exclusivamente à compactação de músicas, notou-se a falta de informações sobre as mesmas incluídas no arquivo. Com isso, foram desenvolvidos os denominados campos ID3 *tag* que contêm informações variadas sobre a música: título, artista, álbum, ano, gênero musical, comentários e trilha do CD a que pertence originalmente. Estes campos-textos de informação não são padrões oficiais, mas têm evoluído e sido adotados como um padrão informal (NILSSON, 2001).

O MP3 é um *codec* de código aberto e disponível para desenvolvimentos proprietários, o que acabou resultando em um variado grupo de *codecs*, que originam arquivos com variadas qualidades de som. Os obstinados pela qualidade do som garantem que a diferença pode ser grande entre um e outro *codec* baseado no MP3 original. Entre os *codecs* MP3 mais conhecidos temos: Xing, Fraunhofer, Blade e LAME.

### 3.1.2.3- Usos, Vantagens e Desvantagens

Em comparação com o formato Windows Media Audio (WMA), o MP3 consegue sons de melhor qualidade em taxas de bits maiores( acima de 90 Kbps), ao passo que o WMA consegue manter qualidades bem mais aceitáveis do que o MP3 quando abaixo de 60 Kbps até 16 Kbps. Outra capacidade do formato MP3, que lhe garante estar no topo da lista dos formatos musicais digitais, é que existe a capacidade de VBR (*variable bit rate*), taxa de bits variável. O MP3 pode ser codificado com várias taxas de bits, mas no modo padrão estas taxas são fixas, ou seja, utilizadas no arquivo todo. Quanto maior a taxa de bits melhor a qualidade do som. Quando codifica-se um arquivo em MP3 com VBR tem-se que taxas baixas (32 Kbps) serão utilizadas em partes da música onde o sinal é fraco ou pobre em complexidade de frequências, e em outros momentos, a taxa de bits subirá até o limite estabelecido pelo usuário no momento da geração do arquivo (192, 320 Kbps). Esta técnica garante uma grande qualidade com menor tamanho ao mesmo tempo. O único inconveniente é que o tempo de codificação é maior (DUNN, 2001).

Na realidade, a atual principal vantagem do MP3 é a sua grande difusão e aceitação pelos usuários no mundo todo, havendo um variado estoque de arquivos de áudio neste formato disperso pelo mundo. Há *encoders*, *players* e *softwares* variados de MP3 para todo tipo de plataforma. Contudo, devido a variantes proprietárias do *codec* MP3, certos *players* não tocam certas variantes de MP3 por incompatibilidade. Veja o Quadro-2 com o resumo das vantagens e desvantagens dos formatos.

Um fato que está fazendo com que muitos servidores WEB estejam sendo configurados para *streaming* de MP3, é o de que a tecnologia Macromedia Flash, que se tornou o padrão de difusão multimídia na Internet ( presente em mais de 85 % dos *browsers* dos usuários), usa o som em formato MP3.

O uso do MP3 pode ser direcionado à atividade geral, como a distribuição de músicas, sonorização de *sites*, de CD's-ROM, *audio-books* e *streaming* de áudio.



### **3.1.3- Liquid Audio**

#### **3.1.3.1-Histórico**

O formato Liquid Audio é um sofisticado sistema de distribuição de músicas pela WEB baseado no sistema Dolby Digital e no *codec* MPEG-2 AAC. Seguiu os passos de seu antecessor A2B (veja adiante o item 3.1.4), mas ao contrário dele, conseguiu o seu objetivo, firmou-se no mercado.

#### **3.1.3.2- Características**

Os arquivos podem ser descarregados diretamente ou escutados por *streaming* . Possuem sistema de marca d'água (*watermarking*) e encriptação para proteção de direitos autorais (Fries, 2000). Um fato um tanto confuso com o formato é de que ele pode vir codificado tanto em Dolby AC-3 quanto na forma AAC, tendo ambos os arquivos a extensão “.LQT” . Mesmo o seu *player* não mostra o tipo de *codec* utilizado (5DDESIGN, 2001).

#### **3.1.3.3- Usos, Vantagens e Desvantagens**

Está entre os fortes concorrentes para se tornar um dos padrões de distribuição e venda de música pela Internet. Trata-se de um formato sólido de características profissionais com garantia de qualidade contra ruídos das mais variadas origens, como “pops” e “glitches”. Este formato pode incluir a arte gráfica do álbum em questão, as letras das músicas, preço, além de outras informações como *links* a locais onde a música ou o álbum podem ser adquiridos. Contudo a instalação de um sistema servidor Liquid Audio não é barata, sendo que a empresa está personalizando o sistema para cada cliente (5DDESIGN, 2001) (Quadro-2).

### **3.1.4- A2B**

#### **3.1.4.1- Histórico**

Formato de distribuição musical desenvolvido pela AT&T baseado no *codec* MPEG-2 AAC. Foi a primeira aplicação comercial de grande porte para o *codec* AAC, sendo lançado em 1997 (5DDESIGN, 2001).

#### **3.1.4.2- Características**

Possui sistema de marca d'água e de encriptação para proteção de direitos autorais, que limita o número de usuários para um arquivo musical e também o número de vezes em que uma música será tocada. Este formato consegue melhor qualidade de som em menores taxas de bits do que o MP3, com compactação de até 20:1(5DDESIGN, 2001).

#### **3.1.4.3- Usos, Vantagens e Desvantagens**

Possui um flexível sistema de licenciamento eletrônico que pode controlar como uma música é utilizada e distribuída. A música codificada com A2B pode incluir a arte gráfica, letra musical, autoria e *links* para o *site* do artista. O fato mais negativo é de que o grande público tende a evitar formatos proprietários altamente fechados e regradados como o A2B. Atualmente (fevereiro de 2003) este formato parece ter sido retirado de circulação e não ter conseguido o objetivo esperado: tornar-se um padrão para distribuição e venda de músicas *on-line* (5DDESIGN, 2001).

### **3.1.5- QuickTime**

#### **3.1.5.1- Histórico**

Formato proprietário desenvolvido pela Apple Computer a partir de 1991. No início foi utilizado somente na plataforma Apple, contudo, posteriormente passou a atuar

também na plataforma Windows (ano de 1992). É um dos 2 principais padrões de vídeo para os usuários caseiros, junto com o formato AVI da empresa Microsoft.

### **3.1.5.2- Características**

Suporta tanto *streaming* de áudio quanto de vídeo. Este formato serve de base para parte do superformato multimídia MPEG-4. O Quicktime é muito mais do que um formato, na realidade é um sistema de gerenciamento multimídia (geração, transporte e execução de mídias) que suporta vários formatos e *codecs* de áudio (*samples* de som e MIDI), animações, 3D, vídeos e até realidade virtual.

### **3.1.5.3- Usos, Vantagens e Desvantagens**

Apesar de também apresentar boa qualidade, o sistema de *streaming* Quicktime foi poucas vezes observado entre as rádios-WEB (ver seção 3.2.2), sendo superado em ocorrência pelos sistemas de *streaming* Windows Media Audio, RealAudio e MP3. Ele não foi criado especificamente como formato para *download* de áudio. Outros formatos como o MP3, WMA, Ogg Vorbis, RealAudio e Liquid Audio já tomaram a ponta por disputa nesta área de mercado. Contudo a empresa Apple, criadora do QuickTime, lançou a versão 6 deste formato, dando suporte ao *codec* MPEG-4 AAC (que tem como base o eficiente *codec* MPEG-2 AAC), assim tentando entrar no promissor mercado de música *on-line* (Quadro-2).

O formato QuickTime tem a filosofia para ser utilizado como um formato multimídia completo tal qual o Real System com a linguagem SMIL e os formatos Flash e Shockwave, mas a sua arquitetura de melhoramento modular o mantém sempre um formato competitivo e seguidor das tendências inovadoras.

### **3.1.6- Macromedia Flash e Shockwave**

#### **3.1.6.1- Histórico**

Estes dois formatos são na realidade duas tecnologias para multimídia na Internet e não apenas formatos de áudio. O formato Shockwave foi lançado em 1995 e era uma forma compactada para a Internet dos arquivos multimídias do *software* Macromedia Director. Logo após, em 1996, era lançado o padrão Flash, uma tecnologia de animação vetorial voltada também à WEB.

#### **3.1.6.2- Características**

Na verdade, o *codec* de áudio que ambas as tecnologias utilizam é o MP3 (MPEG-1 camada-III), mas devido a grande interação do áudio como uma das várias características multimídia destes dois formatos, pode-se de certa forma, considerá-los como uma forma variante do padrão MPEG.

#### **3.1.6.3- Usos, Vantagens e Desvantagens**

A maior vantagem do áudio nestes dois formatos multimídia é justamente a grande capacidade de interatividade e sincronização sonora entre os componentes de uma apresentação multimídia pela Internet.

Os *streamings* de áudio maiores (maiores que 1 minuto ), feitos nas formas Flash e Shockwave, podem ser realizados através da inclusão da tecnologia de RealAudio, que se denomina de RealFlash no caso do Flash. Assim teríamos uma tecnologia híbrida (Quadro-2).

### **3.1.7- Wave**

#### **3.1.7.1- Histórico**

Formato desenvolvido pela Microsoft e IBM para ser o formato de som padrão em microcomputadores com o sistema Windows. Apesar de ser muito criticado como um formato obsoleto, ele permanece até hoje como um dos formatos padrões da Internet. A Microsoft desenvolveu o formato similar RIFF (*interchange File Format*) para contornar alguns problemas de compatibilidade (estruturação interna dos bytes) que existiam com o formato WAVE.

#### **3.1.7.2- Características**

Com o surgimento de novos *codecs*, o formato WAVE pôde sair da redoma do PCM descompactado e agora pode assumir vários tipos de compressão. Entre eles teríamos o  $\mu$ -law, a-law, GSM, IMA ADPCM, MS ADPCM (*codec default*) e até mesmo a compressão MPEG (FRANKE, 2001).

É um formato flexível, podendo a taxa de amostragem, tamanho de amostra e número de canais serem ajustados arbitrariamente.

#### **3.1.7.3- Usos, Vantagens e Desvantagens**

É um padrão universal de som dentro da plataforma Windows, contudo já é lido também por outras plataformas. A grande maioria dos *softwares* de edição de som suporta este formato. Em geral é usado com o *codec* PCM, ficando muito grande o arquivo nesta forma. Este é um formato sem *codec* específico, podendo ser codificado com vários tipos de compressores. Conforme o tipo de compressão, pode haver incompatibilidades entre certos *softwares* de áudio e com outras plataformas como Apple e Unix, por isso, este formato é quase sempre usado na codificação PCM para manter a portabilidade. Salvar um arquivo de som em WAVE (PCM) é o que geralmente se faz nas primeiras etapas de edição de som em um computador pessoal (Quadro-2).

### 3.1.8- AIFF

#### 3.1.8.1- Histórico

O AIFF (*Audio Interchange File Format*) é um formato desenvolvido pela empresa Apple para se tornar o padrão de som dos computadores Macintosh. Acabou sendo utilizado por outras empresas depois, como a Silicon Graphics. O formato .AIFF, o .WAV (windows), o .AU (Unix) e o padrão MIDI foram os primeiros padrões a se estabelecerem na Internet. Mesmo nos dias de hoje (ano de 2003), a quantidade de arquivos no formato AIFF, à disposição na rede mundial, é grande.

#### 3.1.8.2- Características

É um formato bem flexível, dando suporte à configuração arbitrária da taxa de amostragem, tamanho de amostra e número de canais (estéreo e mono). Existe o padrão compactado deste formato (AIFFC) que se utiliza de *codecs* proprietários (MACE) e também gerais como o  $\mu$ -law e IMA ADPCM (FRANKE, 2001).

O AIFFC foi tido como o sucessor do AIFF, visto que tinha todas as características dele e outras melhorias mais. Contudo, isto não está se concretizando completamente, pois muitos *softwares* de áudio ainda se mostram incompatíveis com o AIFFC, principalmente na forma compactada.

O AIFF inclui cabeçalho de informação no arquivo que contém: nome do arquivo, taxa de amostragem, tamanho do arquivo e resolução em bits. Também suporta resolução de 24 e 32 bits (HASS, 2002).

#### 3.1.8.3- Usos, Vantagens e Desvantagens

Uma das grandes vantagens do formato AIFF é a sua capacidade multiplataforma acentuada, sendo reconhecido também por todos os *browsers* modernos. É formato padrão da maioria dos *softwares* de som. Apresenta características semelhantes ao formato WAVE (suporta vários *codecs* e taxas de amostragem), só que é mais direcionado para a plataforma Apple (Quadro-2).

É formato de uso geral, contudo, deve-se direcioná-lo mais para a plataforma de CD-ROM e para a edição de áudio, visto que formatos diferentes já o superaram em outras áreas de uso.

### **3.1.9- Sun Audio (AU)**

#### **3.1.9.1- Histórico**

Formato nativo do Unix, existia em duas versões, uma criada pela empresa Next (que suportava mais de 20 tipos de compressão) e outra pela Sun Microsystems (que suporta 3 *codecs*). Apenas a versão da Sun perdurou e acabou por ser tornar em um dos formatos mais difundidos em toda a Internet (HENRY, 2001).

#### **3.1.9.2- Características**

O formato Sun Audio suporta 3 tipos de *codecs*: PCM, A-law e mu-law. A extensão é .AU e é o formato de áudio padrão para a linguagem de programação Java. O *codec* mu-law é o que se tornou padrão para a Internet neste formato (PATTERSON & MELCHER, 1998).

#### **3.1.9.3- Usos, Vantagens e Desvantagens**

Suportado em praticamente todas as plataformas. A qualidade do som para os dias de hoje já não é mais aceitável no que se refere quando na forma de música, pois o tamanho de amostra é de 8 bits (*codec* mu-law). Os *browsers* Netscape e Internet Explorer apresentam capacidade de suporte embutida a este formato.

Este formato de som, devido as suas características, pode ser usado como clipes curtos de som (menos de 30 segundos), fundo musical, efeitos de som e clipes de voz, sendo estes últimos os mais indicados (Quadro-2). Contudo, a tendência é de se evitar o uso deste formato nestas atividades, pois existem outros formatos mais adequados para tais finalidades.

### **3.1.10- RealAudio ( RealSystem G2)**

#### **3.1.10.1- Histórico**

Formato criado pela empresa RealNetworks (antes Progressive Networks) em 1995. No início as transmissões eram rudimentares e sem *streaming* em tempo real (ao vivo). No entanto, com as novas implementações, o formato RealAudio tornou-se um padrão firme na Internet. Foi o formato de *streaming* pioneiro na Internet. O sistema RealSystem G2 é o primeiro sistema a implementar o padrão RTSP (*Real Time Streaming Protocol*) para *streaming* em tempo real. Dentre as várias versões, aquela que se tornou um marco foi a versão 6.0 que foi a que realmente se mostrou efetiva em recursos, ficando conhecida como RealAudio G2 (*“second generation”*).

#### **3.1.10.2- Características**

O RealAudio consiste de três componentes de *software*: um *player/plug-in* para o *browser*, um *encoder*, e um servidor. Tem a extensão .RA para áudio, podendo vir também como .RM, sendo comumente vídeo com esta extensão. O formato RealAudio se compõe na realidade de várias versões do mesmo arquivo codificadas com qualidades diferentes para ocorrer uma negociação de largura de banda entre o cliente (computador do usuário) e o servidor de arquivos (Real Server) no momento da conexão. Assim, conexões dos mais variados tipos (28,8; 33,6; 56; 128 Kbps e outras) podem ter arquivos melhorados para as suas características (THB.MUSIC SERVICES, 2001).

A partir da sua versão G2, passou-se a contar com uma tecnologia que vai além de simples negociação de banda de rede, uma tecnologia que pode mudar dinamicamente a taxa de bits do *streaming* para acomodar variações na largura de banda da rede ( BEGGS & THEDE, 2001).

É importante se observar que, apesar de ficarem armazenados em servidores WEB, um arquivo Real Audio, em grande parte das vezes não poderá ser gravado no computador do internauta. Isso porque a maioria dos *sites* que utilizam esta tecnologia o fazem de modo que seus servidores só permitam o *streaming* dos



arquivos via Real Player. O *download* dos arquivos, que possibilitaria que o mesmo fosse salvo em disco, é vedado aos internautas. Até meados do ano 2000, 85% das rádios *on-line* utilizavam a tecnologia Real Audio para transmitir suas programações. No entanto, a entrada no mercado do Windows Media (Microsoft) vem abalando a liderança da Real Networks (CENTRAL DA MÚSICA, 2002a).

Na realidade, este sistema de *streaming* vai além da tecnologia de apenas áudio, pois cobre e suporta outros tipos de mídias: como RealText (textos) e RealPix (imagens). Além disso, tem agora o suporte para SMIL(*Synchronized Multimedia Integration Language*), uma linguagem no estilo da HTML para gerenciar a sincronização da apresentação de várias mídias ao mesmo tempo. A arquitetura aberta do RealSystem G2 (que inclui o RealAudio) permite adicionar novas tecnologias e dados, como MIDI, VRML e MPEG (TERRAVISTA, 2002).

### 3.1.10.3- Usos, Vantagens e Desvantagens

Uma das grandes vantagens é a capacidade de que o sistema cliente/servidor RealAudio tem de disponibilizar várias qualidades de *streaming* para um mesmo arquivo conforme a capacidade de conexão do usuário. É a chamada tecnologia **SureStream** que permite ajustar dinamicamente a taxa de bits de modo a ajustar-se à largura de banda disponível, eliminando o *rebuffering* (Quadro-2).

Este formato é impróprio para apresentações interativas multimídia com respostas sonoras a ações do mouse ou teclado, pois é necessário alguns segundos para a análise pelo servidor do pedido do usuário. Ações interativas sonoras são apropriadas para os formatos Flash, Shockwave e Beatnik, requerendo uma completa solução cliente, muito mais do que uma solução servidor.

O formato RealAudio tem larga aceitação em vários usos pela WEB: *audio-books*, *streaming* de rádios, arquivos de sons históricos e outros arquivos de informações sonoras como: aulas, notícias e comentários.

### **3.1.11- Windows Media**

#### **3.1.11.1- Histórico**

O formato Windows Media Audio, criado pela empresa Microsoft, chegou ao mercado em 1998 como uma tecnologia de *Internet broadcasting* (difusão de mídias pela WEB). Depois de várias atualizações a Microsoft passa a expor que o seu formato é superior ao formato MP3, sendo que a codificação de 64 Kbps teria a mesma qualidade do que o MP3 de 128 Kbps (THOMASON, 2001).

#### **3.1.11.2- Características**

O WMA em *streaming* assume a extensão ".ASF". O formato ASF é na realidade uma forma de empacotar vários formatos de mídias, e até mesmo comandos para gerenciamento de mídias, sendo algo de certa forma na filosofia de Flash, Shockwave e Real SMIL.

Atualmente é o formato que mais cresce devido a várias estratégias mercadológicas da empresa Microsoft, além de grandes avanços tecnológicos como o VBR e o som surround 5.1 (WMA versão 9).

#### **3.1.11.3- Usos, Vantagens e Desvantagens**

Uma das vantagens do WMA sobre o MP3, é que o WMA consegue melhor qualidade de som com taxas de bit mais baixas (abaixo de 90 Kbps, até mesmo 16 Kbps), ao passo que em maiores taxas de bit o MP3 faz um trabalho melhor. O WMA não possuía capacidade VBR como o MP3 (DUNN, 2001). Entretanto hoje em dia, com a versão 9, a Microsoft introduziu o VBR e até o som multicanal.

Este formato atualmente, junto com o RealAudio e o MP3, é um grande padrão para *streaming* de rádios-WEB, transmissão de eventos jornalísticos, venda de *audio-books* e de músicas pela Internet. Muitos dispositivos portáteis para ouvir o formato MP3 já começam também a suportar este formato (inclusive DVD *players*) (Quadro-2).

É um formato de uso geral para a Internet e tem sido atualizado de tempos em tempos com várias melhorias. A empresa Microsoft investe fortemente neste formato para que ele venha a se tornar em um dos grandes padrões de áudio para a WEB.

### 3.1.12- VQF

#### 3.1.12.1- Histórico

O "*Transform-domain Weighted Interleave Vector Quantization*", ou mais resumidamente "*Vector Quantization Format*" (VQF), é um formato desenvolvido posteriormente ao MP3 (padrão MPEG) pelo **NTT Human Interface Laboratories** do Japão e implementado comercialmente pela empresa Yamaha a partir do ano de 2000, que além de produzir carros, motos e outras tecnologias de veículos, também está presente como fabricante de teclados e sintetizadores no mercado musical. Trata-se de uma tentativa de estabelecer um padrão de formato digital de som, visando fortificar o controle do mercado já conquistado (VQF.COM, 2001).

Princípios deste formato podem ser utilizados no projeto de desenvolvimento do multifuncional padrão MPEG-4 (BOUVIGNE, 2001).

Atualmente (janeiro de 2003), o *site* que dá as especificações deste formato, informa que devido a estratégias de mercado, o formato está suspenso, podendo ser retomada a sua implementação no futuro, conforme novas implementações de versões sejam adicionadas. Contudo, dada as suas características bem diferentes em relação aos outros *codecs*, este formato deve ser mantido em pauta.

#### 3.1.12.2- Características

Este formato tem um poder de compactação de 25 a 35% maior que o MP3 e isto em uma taxa de bits menor (96 kbps). A opinião de muitos é de que o arquivo em 96 Kbps e 44,1 Khz de um VQF soa bem melhor que um MP3 de 128 Kbps e 44,1 Khz. As opiniões de quanto melhor um VQF soa em relação a um MP3 podem variar. Temos dois grandes fatores influenciando estas opiniões: a sensibilidade que varia de pessoa à pessoa e o problema de que os *encoders* de MP3 podem ter variações

conforme a empresa que os implementa. Os *encoders* de VQF tem basicamente o mesmo algoritmo. O formato VQF foi desenvolvido direcionado para baixas taxas de bits (menores que 128 Kbps). Caso seja necessário o uso de qualidades mais elevadas (256, 320 kbps), então o uso do MP3 será o indicado. Neste caso, tem-se que tanto o Windows Media (WMA) e o VQF são indicados para baixas taxas de bits, não conseguindo superar a qualidade do MP3 em taxas mais altas (VQF.COM, 2001).

O VQF usa alguns princípios de algoritmo também utilizados no MP3 e no MPEG-2 AAC, contudo o princípio base é diferente. Os bits de som não são codificados individualmente, mas são unidos em padrões vetoriais, sendo que estes padrões formados são comparados com padrões pré-estabelecidos numerados. Achando o padrão pré-estabelecido que mais se aproxima do formato, então apenas o número de identificação deste padrão é codificado.

### **3.1.12.3- Usos, Vantagens e Desvantagens**

Existem poucas músicas disponíveis neste formato. A codificação é mais lenta e exige melhores processadores. Ao contrário do MP3, que é um formato desenvolvido por uma associação sem fins lucrativos, o MPEG (licenciamento pago, contudo não é caro), o VQF é desenvolvido por uma empresa comercial e sua distribuição para usuários é gratuita (*players*, *softwares* de suporte, *encoders*). Isto pode parecer um tanto paradoxal. De certa forma, quem não visa lucro vende e quem visa lucro distribui gratuitamente (VQF.COM, 2001). Outro fato, é que a Yamaha nunca promoveu o formato, e tem licenças pesadas para alguém que queira trabalhar comercialmente com ele, resultando em pouco desenvolvimento exterior.

### **3.1.13- Midi**

#### **3.1.13.1- Histórico**

O MIDI nasceu como um método de comunicação entre instrumentos musicais digitais e computadores no princípio dos anos 1980, utilizando-se da tecnologia da sintetização digital de sons. O fato de ter se tornado um padrão na rede mundial

praticamente deve-se ao seu pequeno tamanho de arquivo. Apesar das suas limitações, o MIDI ainda hoje é bastante utilizado em páginas na Internet. Tem-se que ter em mente que este formato não foi desenvolvido para a WEB e sim para a música e efeitos sonoros.

### **3.1.13.2- Características**

Basicamente em um arquivo MIDI não temos nenhum evento sonoro ("*sample*") arquivado nele como nos arquivos de som reais, tal como o caso dos arquivos .WAV (padrão Windows) e .AIFF (padrão Apple), mas tem-se apenas instruções para um sintetizador de som gerar as notas musicais em uma seqüência determinada e com isso emitir uma composição musical na sua saída. Poderia-se dizer que o que está no arquivo MIDI é uma partitura digital que pode até ser impressa como tal. O som em si não existe, o que existe são instruções para tocá-lo (SCHERER, 2001).

O tamanho de um arquivo MIDI é determinado mais pela complexidade da música do que pela sua duração. Pode-se entender por complexidade da música todos os comandos MIDI necessários para gerar tal música, não se deve esquecer que nem todo comando MIDI se refere sempre a tocar ou não uma determinada nota musical, mas temos comandos de controle, efeitos e outros mais (ANDERSSON, 2001).

### **3.1.13.3- Usos, Vantagens e Desvantagens**

A grande vantagem do arquivo MIDI é o seu pequeno tamanho que pode chegar a ser facilmente 100 vezes menor que um arquivo MP3, que por sua vez é cerca de 10 vezes menor que um arquivo WAVE (descompactado PCM). A outra grande vantagem do arquivo MIDI é que por ser já uma partitura em si, com as instruções para gerar as notas musicais já prontas, ele pode ser editado e alterado com *softwares* de edição e composição MIDI, os chamados sequenciadores, como Cakewalk Pro, Sonar, Steinberg CubBase VST, Steinberg Nuendo e E-Magic Logic Audio, com grande facilidade. Com *softwares* como estes, o poder de edição é completo e os instrumentos musicais estão separados em trilhas diferentes. Os instrumentos de cada trilha podem ser alterados à vontade do editor de áudio, sendo que, por exemplo, uma clarineta pode se transformar em um violino ou trombone. O

número de instrumentos disponíveis é muito grande, mas deve-se ter em mente que a qualidade da saída do som dependerá da qualidade do sintetizador de som que se usa, seja o sintetizador de qualquer dispositivo MIDI ou da placa de som de um computador. Tem-se basicamente 2 tipos de sintetizadores em placas de som, o sintetizador FM que literalmente cria os sons gerados e o sintetizador *wave-table* que possui uma memória interna com uma tabela grande de *samples* de sons dos instrumentos em si. Sendo assim, o sintetizador *wave-table* é sempre melhor em relação ao FM pois utiliza uma tabela de sons reais e não sons artificiais, porém, a qualidade e o preço dos sintetizadores variam bastante (SCHERER, 2001).

Um dos maiores problemas do formato MIDI é que apenas sons musicais podem ser armazenados, sons orgânicos ou naturais não podem ser adicionados, como a voz humana, um latido de um cão e o barulho de um relâmpago. Pode-se compor amostras reais de sons com sons MIDI, porém, o resultado final deverá ser exportado para um formato não-MIDI (*wave-PCM* por exemplo). Com a evolução dos computadores e instrumentos musicais, o formato MIDI foi sem dúvida um dos que mais evoluiu. A interação entre instrumentos MIDI e computadores é grande. O fato de apenas tocar o teclado e observar as notas serem geradas imediatamente como partituras na tela do computador através de *softwares* especializados faz a satisfação dos músicos atuais (ZHANG, 2001). A facilidade de criação de sons MIDI pode ser muito grande, pois existem vários *softwares* conhecidos como conversores *wave-to-MIDI*, ou seja, que convertem sons reais (*samples* ou amostras) para comandos MIDI. O *software* mais conceituado e respeitado destes é o *Autoscore*. Com este *software* pode-se cantarolar qualquer melodia em um microfone sendo a voz posteriormente convertida em comandos MIDI, que podem assumir qualquer instrumento musical disponível naquele momento. O único problema é que este arquivo MIDI será tocado com apenas um instrumento de cada vez, ou seja, apenas uma trilha de comando MIDI. Este é o chamado arquivo monofônico, ao contrário dos arquivos mais modernos que possuem 32, 64 e até mais instrumentos ao mesmo tempo (o resultado assim fica pouco profissional).

Atualmente os arquivos MIDI são utilizados mais em páginas pessoais e como fundo musical, contudo a utilização como divertimento no formato MIDI Karaokê tem se difundido pela Internet. Os formatos “karaokê” não passam de arquivos MIDI proprietários que são tocados por *players* específicos. Estes arquivos são executados

em um aparelho de karaokê virtual no computador com imagens e a letra da música aparecendo no monitor exatamente como ocorre nos aparelhos convencionais de karaokê. Vários *softwares* para edição e execução de MIDI Karaokê estão disponíveis: RealOrche, Midioke, EasyKAR, STAR3, WinKaraoke Producer e Karaoke Producer (HELENIO, 2001).

Uma ressalva a ser feita com o padrão MIDI, é de que ele também tem formas variantes (GM, XG, GS), podendo haver certa incompatibilidade entre elas (Quadro-2).

### **3.1.14- MOD**

#### **3.1.14.1- Histórico**

Os arquivos MOD originaram-se nos computadores Commodore Amiga (estes computadores não são mais produzidos hoje em dia), contudo, devido a sua flexibilidade e ao elevado número de arquivos disponíveis atualmente, os MOD *players* estão disponíveis para uma grande variedade de plataformas (BAGWELL, 2001a).

#### **3.1.14.2- Características**

Os arquivos MOD são arquivos de som (quase sempre música) contendo duas partes:

- a) um banco de amostras digitalizadas;
- b) comandos descrevendo como e quando estas amostras são tocadas.

As amostras em um arquivo MOD são RAW (descompactadas) e com 8 bits de resolução. Pode haver até 31 amostras em um arquivo, cada uma delas com um tamanho de até 128 Kb. Formatos antigos permitiam apenas 15 amostras por arquivo. Não há frequência de amostragem padronizada para estas amostras. A informação do seqüenciamento do som em um arquivo MOD contém 4 trilhas que descrevem quais, quando, por quanto tempo e com que frequência as amostras devem ser

tocadas. Isto significa que um arquivo MOD pode ter até 31 sons digitalizados de instrumentos distintos com até 4 instrumentos tocando ao mesmo tempo em um dado momento. Isto permite uma grande variedade de possibilidades de melodias, sendo possível a inclusão de amostras de voz ou a criação de instrumentos pelo próprio usuário (com o uso de um *sampler* apropriado). Esta habilidade de usar suas próprias amostras como instrumentos é uma flexibilidade que outros formatos musicais não compartilham, e este é um dos motivos que faz o arquivo MOD ainda tão popular, numeroso e variado (BAGWELL, 2001a).

### **3.1.14.3- Usos, Vantagens e Desvantagens**

Um dos grandes problemas do padrão MOD é que existem muitas variantes com capacidades de efeitos diferentes, muitas incompatíveis entre si. Isto limita o poder de criação do editor de som dentro das capacidades de cada variante. Estas variantes estão intimamente ligadas ao *software* ("tracker") que as geraram. O *player* que toca um determinado tipo de MOD pode não tocar outro arquivo variante. O tamanho do arquivo MOD está mais ligado à complexidade da composição do som e da quantidade de *samples* incluída nele do que com o tempo de duração da própria música ou composição em si, tal qual o arquivo MIDI. Quando as condições de um *player* correto e placa de som compatível são respeitadas, o resultado ouvido é uma reprodução fiel em diferentes computadores. Estes arquivos podem atingir grandes tamanhos mesmo em composições curtas, o que em geral é meio perturbador para quem está acostumado a lidar com a edição de arquivos de som tradicionais como WAVE, AIFF e MP3 (ANDERSSON, 2001) (ver Quadro-2).

### **3.1.15- Beatnik**

#### **3.1.15.1- Histórico**

Formato híbrido de som (comandos padrão MIDI + amostras de som real) semelhante de certa forma ao padrão MOD. Estas amostras que estão inclusas em um arquivo Beatnik podem estar codificadas em MP3. Este formato foi desenvolvido



pela empresa Headspace (atualmente Beatnik), fundada em 1996, para suprir a parte sonora de um *website*, dando-lhe interatividade e aprimorando o *design*. Mais de 7 milhões de usuários na Internet já utilizam o *player* Beatnik.

### 3.1.15.2- Características

Possui a extensão RMF (*Rich Music Format*). Para a sua execução é necessário a presença de um *plug-in* ou *player* Beatnik e também do suporte *JavaScript Music Object*, que é uma biblioteca em Javascript que permite que o *player* possa comandar o som interativo da página WEB (ADOBE SYSTEMS, 2001).

O processo da tecnologia Beatnik funciona basicamente assim: a música e os comandos de controles de som que uma pessoa coloca em sua página WEB, na forma de código HTML e de Javascript, são baixados ao computador do usuário onde o *browser* lê os comandos (HTML e Javascript), estes comandos então se conectam com o Beatnik Player via a biblioteca “Music Object”, que fornece um grande conjunto de comandos para as ações interativas do usuário. Assim então o arquivo RMF é tocado conforme são disparadas as ações interativas do usuário: clicar, arrastar, mudar de página e outras ações mais (Figura-19) (BEATNIK, 2001a).

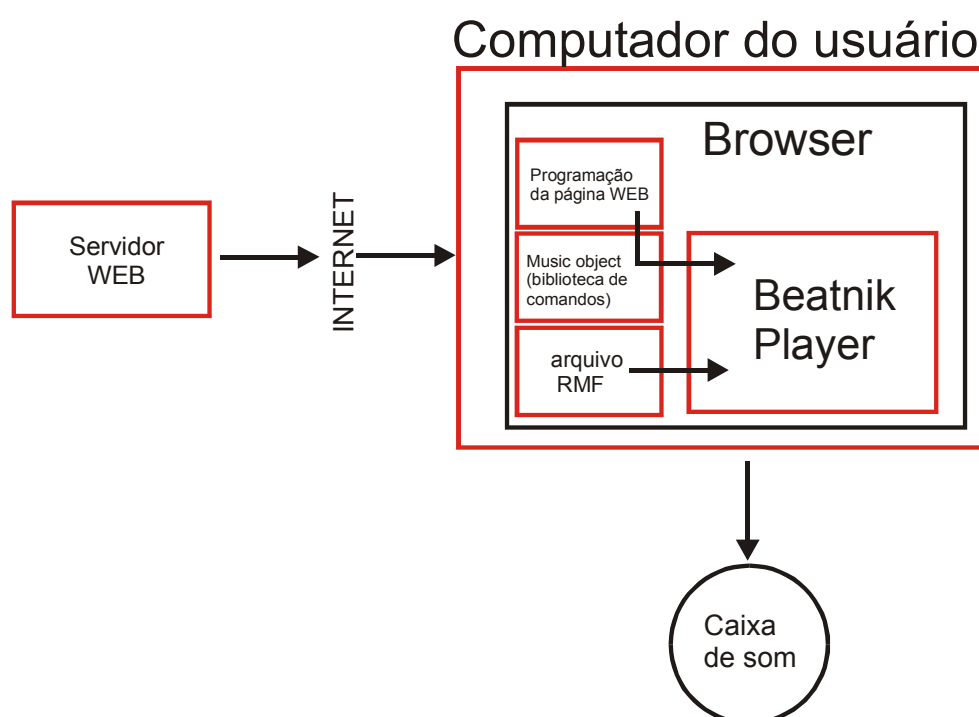


Figura 19 - O processo durante a execução de um arquivo Beatnik.  
 Fonte: ([http://www.beatnik.com/software/tutorials/beatnik\\_basics.html](http://www.beatnik.com/software/tutorials/beatnik_basics.html)).

Apresenta sistema de proteção de direitos autorais. Não se trata de um formato *streaming*, uma vez baixado da rede e começado a tocar o arquivo, o som é constante e nítido.

Os arquivos RMF são utilizados para sonorizar botões, barras de navegação, mudanças de páginas e outras ações do usuário dentro de um *site*. Até mesmo gerações de som em tempo real como mudança de tom e volume conforme as ações do *mouse* são possíveis. Não é exigido nenhum *software* do lado servidor, toda a ação é realizada no lado cliente pelo *player* Beatnik (BEATNIK, 2001b).

### **3.1.15.3- Usos, Vantagens e Desvantagens**

O formato Beatnik introduziu um novo conceito de som na Internet: qualidade, pequeno tamanho de arquivo e principalmente interatividade. Ao contrário do outro formato híbrido, o MOD, este é um formato especialmente direcionado para a WEB.

Uma das grandes vantagens do formato Beatnik é que ele é multiplataforma e qualquer som dele soará do mesmo modo em qualquer uma delas, mesmo sons compostos de comandos MIDI ou amostras de sons reais. O som MIDI puro soa de modo diferente conforme o tipo de sintetizador instalado na placa de som do computador ou no instrumento musical. No caso do formato Beatnik a síntese é sempre feita pelo *player* Beatnik que contém todas as informações necessárias para uma reprodução fiel em várias plataformas operacionais (Quadro-2).

### **3.1.16- Crescendo**

#### **3.1.16.1- Histórico**

Tecnologia desenvolvida pela empresa Liveupdate para ouvir o formato MIDI pela WEB. Apesar de ser um formato muito antigo, o MIDI ainda encontra lugar na WEB, dando margens a implementações de tecnologias baseadas nele como são Crescendo e Beatnik.

### 3.1.16.2- Características

"Crescendo" não é um formato, mas sim, uma tecnologia de *plug-in* e controle *ActiveX* que torna possível o *streaming* de arquivos MIDI. Para gerar o *streaming* de arquivos MIDI é necessário uma chave lógica que cria um diretório para *streaming* no servidor do usuário. É uma espécie de parte servidor do sistema Crescendo denominada Streamsite (THB.MUSIC SERVICES, 2001).

### 3.1.16.3- Usos, Vantagens e Desvantagens

A função de *streaming* tem ainda alguns problemas (dados do ano 2000). Muitos arquivos MIDI precisam alguns ajustes para a completa transferência via *streaming* (Crescendo corta alguma parte no início e final dos arquivos MIDI comuns) (THB.MUSIC SERVICES, 2001) (Quadro-2).

Com a tecnologia Crescendo instalada em um *browser* pode-se escutar praticamente qualquer tipo de arquivo MIDI.

## 3.1.17- Ogg Vorbis

### 3.1.17.1- Histórico

Formato de áudio digital de código aberto, criado por um grupo de programadores espalhados por todo o mundo e centralizados pela instituição chamada **Xiph.Org**. Este formato começou em 1993 como parte de um grande projeto sobre multimídia na rede denominado Projeto Ogg. O nome singular vem de "Ogg" um tipo de estratégia em um jogo chamado Netrek e "Vorbis" é um personagem do romance "Pequenos Deuses" de Terry Pratchett. Tal como o *codec* "DIVX" de vídeo já se pode perceber pelo seu nome um certo elemento "*hacker-nerd*" na origem do Ogg Vorbis (THOMASON, 2001).

### 3.1.17.2- Características

Além de ser de código aberto, o formato gera arquivos de áudio digital 10 a 15% menores que um arquivo gerado por um *codec* MP3, mantendo a mesma qualidade. Muitos apostam nele como o mais forte candidato a sucessor do MP3. O *codec* Ogg Vorbis (.Ogg) só gera arquivos em VBR (taxa de bit variável), o que significa na prática, que músicas mais lentas geram arquivos menores e músicas complexas e mais rápidas, arquivos maiores (CENTRAL DA MÚSICA, 2002b).

O formato Ogg tem capacidade de *streaming* e compressão multicanal como o formato MPEG-2 AAC (LITEXMEDIA, 2003) (Quadro-2).

### 3.1.17.3- Usos, Vantagens e Desvantagens

O Ogg é um formato ainda em desenvolvimento, e não traz, por exemplo, *tags* (conhecidas como "ID3" no MP3) que permitem informar dados sobre as músicas, porém, isto está em implementação. Ele ainda não é muito difundido e conhecido na Internet, o que torna uma tarefa difícil encontrar músicas disponíveis neste formato em bons *sites* (o *site* oficial do formato é uma exceção).

Mas a grande vantagem deste formato é o fato do Ogg Vorbis ter sido criado pela comunidade adepta do código aberto. Assim, qualquer programador de qualquer lugar no mundo pode ajudar no seu desenvolvimento, implementando assim cada vez mais novas melhorias, exatamente como ocorre com o sistema operacional Linux (CENTRAL DA MÚSICA, 2002b). Contudo, isto pode dificultar o estabelecimento de um padrão.

No geral, pode-se dizer que este poderá ser um formato promissor devido ao fato de ser completamente aberto e gratuito, além de possivelmente eficiente, porém, ainda não se mostra um produto acabado e por completo utilizável.

O Ogg Vorbis é, tal qual o MP3, direcionado para o uso geral: *audio-books*, músicas, *streaming* e como fonte de informação em áudio de maneiras variadas.

### **3.1.18 - QDX (QDesign)**

#### **3.1.18.1 – Histórico**

Formato proprietário desenvolvido pela empresa QDesign para áudio exclusivo para Internet, tanto para baixar o arquivo, quanto para *streaming*.

#### **3.1.18.2 – Características**

Basicamente o formato QDX , segundo informações do próprio desenvolvedor deste formato, tem maior qualidade e menor tamanho que o MP3. É necessário apenas um arquivo fonte de 128 Kbps, e não várias codificações, como outros formatos para a distribuição via *streaming* pela rede (qualquer que seja o tipo dela). Através de um sistema de redimensionamento dinâmico único (na mesma filosofia do RealAudio), este arquivo de 128 Kbps se adapta aos vários tipos de redes, larguras de bandas e velocidades de conexão. Tudo isto é feito em tempo real. Tem sistema de proteção de direitos autorais para distribuição de música, o que parece ser um de seus objetivos (QDESIGN, 2001) (Quadro-2).

#### **3.1.18.3 - Usos, Vantagens e Desvantagens**

Segundo o próprio fabricante, o grande diferencial do QDX é a sua fácil codificação que é totalmente transparente ao usuário, não exigindo conhecimentos para configuração do *encoder* e também pelo seu sistema de redimensionamento dinâmico adaptando-se assim para qualquer tipo de situação de rede ou para qualquer dispositivo final, seja um computador *desktop*, um PDA ou até um telefone celular com sua baixa taxa de bits (8 Kbps) (QDESIGN, 2001). Contudo, o formato ainda é pouco testado pelo grande público, sendo as informações fornecidas pelo próprio desenvolvedor, necessitando assim de testes mais abrangentes e em grande escala.

O *codec* QDX foi escrito em linguagem "C" e está projetado para ser suportado em vários sistemas operacionais (multiplataforma). Também tem capacidades para

futuras melhorias já previstas: suporte a "*metadata*" como o ID3 *tag* do MP3 e suporte para áudio DVD (QDX pode codificar áudio de 24 bits até 192 kHz e tem suporte para multicanais) (TERRAVISTA, 2002).

### **3.1.19- Os Formatos Menores e Menos Conhecidos**

A quantidade de formatos de arquivo de áudio digital é grande (ver lista em ANEXO), talvez o número real chegue a casa das centenas. Pode-se incluir aqui os formatos não-Internet como Dolby, Sony, THX e DTS (utilizados em cinema, DVD e *home theater*). A grande maioria não passa de formatos adaptados licenciados junto ao criador e com a inclusão de dispositivos proprietários. Não é raro o aparecimento de formatos maquiados, isto é, copiados. O desvio de direitos autorais na área de codificação de áudio parece não ser fato raro.

O Grupo "*Global Music Outlet*" usa o termo MP4 (o que quase sempre induz a um erro, pois não é a evolução do MP3) para descrever o seu sistema de distribuição digital de música proprietário. É baseado em uma versão melhorada do MPEG AAC e inclui um *player* embutido (cada música é um arquivo .EXE). Outras informações como *links* ao *site* do artista podem ser incluídas no arquivo (FRIES, 2000).

O PlayJ (.plj) é um outro formato de áudio digital desenvolvido pela Everad, Inc., com praticamente a mesma qualidade e tamanho de um arquivo MP3 e especialmente codificado. O formato surgiu para ser uma nova forma de distribuir músicas gratuitamente, sem no entanto desrespeitar os direitos autorais. Isso é possível graças ao PlayJ *Player*, que enquanto toca uma música neste formato, exibe um anúncio pago por um patrocinador qualquer (é a chamada política *adware* de *software*). Mais de 70 selos musicais concordaram em distribuir música neste formato, e dezenas de milhares de músicas que eram somente disponíveis em CD são agora gratuitas através de *download* neste formato (CENTRAL DA MÚSICA, 2002c).

O formato Emblaze Audio (EA) foi criado pelo grupo GEO. É um formato com compressão que é tocado por um aplicativo Java. Alguns afirmam que este é o primeiro formato de áudio especificamente direcionado à Internet. Qualquer *browser*

que suporta Java poderá tocar um arquivo .EA sem *plug-ins* adicionais (PROJECTCOOL, 2001).

Além destes formatos menores e menos conhecidos de cunho proprietário e específicos descritos aqui, existem outros mais que não são citados por estarem além do escopo deste trabalho (veja lista em Anexo).

### **3.1.20- Quadro Geral Comparativo dos Principais Formatos de Arquivos de Som em Uso Detectados na WEB**

No Quadro-2 pode-se observar um resumo com as principais características dos formatos detectados neste trabalho (ver seção 3.1). Em alguns casos não se pôde detectar certas características de alguns formatos, pois estas informações não se mostraram disponíveis, e porque também em alguns casos tal dado não era compatível com o formato. Exemplo típico é a determinação de taxa de bits em um arquivo MIDI, pois o som é sintetizado.

Quadro 2- Quadro comparativo resumido com as características dos formatos de áudio encontrados durante este trabalho. Os formatos VQF e A2B não foram inclusos, pois foram retirados de uso do mercado pelos seus criadores.

	CRIADORES	DIREITOS AUTORAIS	STREAMING	TIPO DE SOM	TAXA DE BITS	ÁREA DE USO	VANTAGENS	DESVANTAGENS	FORMAS VARIANTES
<b>MP3</b>	Grupo MPEG	Licenciado e de código aberto	Sim	Sample	32-320 Kbps	Geral, música, rádio, audio-book	Uso comum, boa qualidade, boa compactação, VBR	Variantes c/ incompatibilidade	Sim
<b>MPEG-2 AAC</b>	Grupo MPEG	Licenciado p/ uso profissional e de código aberto	Sim	Sample	8-320 Kbps	Geral de uso profissional: rádio, DVD, música	Qualidade excelente, multicanal. E o codec c/ mais recursos.	Licenciamento caro	Vários formatos baseados nele
<b>LIQUID AUDIO</b>	Liquid Audio	Marca d'água, encriptação	Sim	Sample	Similar ao MPEG-2 AAC	Venda de música pela WEB	Qualidade de som, proteção de direitos bem difundido	O codec pode variar	Não
<b>QUICKTIME</b>	Apple	Com sistema de proteção	Sim	Sample	Similar ao MPEG-2 AAC	Rádio-WEB, venda de música, audio-book.	Muito conhecido, boa qualidade	É um formato multimídia não tendo codec definido	Não
<b>FLASH SHOCKWAVE</b>	Macromedia	Com sistema de proteção	Pseudostreaming e streaming real com plug-in (Real/Audio)	Sample	8-160 Kbps	Sonorização interativa de sites, CD-ROM's, jogos	Padrão mundial, alta compactação	É um formato multimídia muito compactado, com sem qualidade para música	Não
<b>WAVE AIFF</b>	Microsoft e Apple	Sem sistema de proteção	Não	Sample	Definição por conta do editor do som	Edição de som e usos gerais comuns: CD-ROM e sites	Padrão mundial multipataforma	Formato antigo que tem vários codecs, podendo haver incompatibilidade	Sim (vários codecs)
<b>SUN-ÁUDIO</b>	Sun Microsystems	Sem sistema de proteção	Não	Sample	Definição por conta do editor do som	Sonorização interativa e arquivamento de som em sites	Padrão mundial multipataforma	Formato antigo de baixa qualidade de som	Sim (3 codecs)
<b>REALAUDIO</b>	Real Networks	Com sistema de proteção	Sim	Sample	Ajustável dinamicamente	Rádio-WEB, audio-book, aulas, notícias, bancos de sons históricos	Padrão mundial multipataforma, com taxa de bits dinâmica	Sem qualidade musical boa de som, porém inovações já são previstas	Não
<b>WMA</b>	Microsoft	Com ou sem sistema de proteção, conforme o criador do arquivo	Sim	Sample	40-355 Kbps (VBR) 48-192 Kbps (CBR)	Rádio-WEB, audio-book, aulas, notícias, bancos de sons históricos, música	Padrão, grande qualidade, multicanal, VBR, negociação de banda, suporte de grande empresa	Futuramente, poderá haver limitações de uso	Não
<b>MIDI</b>	Tratado internacional	Aparentemente sem sistema de proteção	Sim (Crescendo)	Sintético		Sonorização de sites, músicas, karaokê	Tamanho mínimo de arquivo, facilidade de edição	O som varia conforme o tipo de sintetizador, ausência de sons naturais	Sim
<b>MOD</b>	Commodore	Aparentemente sem sistema de proteção	Não	Híbrido	Dado não disponível	Composição musical	Facilidade para composição musical, grande número de adeptos, com sons naturais	O arquivo pode ser grande, variantes incompatíveis	Sim
<b>BEATNIK</b>	Beatnik	Com sistema de proteção	Não	Híbrido	Dado não disponível	Sonorização interativa de sites	Som interativo, arquivo pequeno	Edição do arquivo RMF exige treinamento	Não
<b>CRESCENDO</b>	Liveupdate	Aparentemente sem sistema de proteção	Sim	Sintético		Sonorização de sites, arquivamento de músicas	Streaming de formato MIDI	Com linhas rápidas não tem grande utilidade; presença de bugs.	Não
<b>OGG VORBIS</b>	Xiph.Org	Sem sistema de proteção, código aberto e gratuito	Sim	Sample	Similar ao MP3	Rádio-WEB, audio-book, aulas, notícias, bancos de sons históricos, música	Formato totalmente aberto a programadores, VBR, streaming, multicanal	Pouco difundido, tende a ser um formato com bugs, devido à implementação de código aberto	De certa forma sim (código aberto)
<b>QDX</b>	Qdesign	Com sistema de proteção	Sim	Sample	128 Kbps	Streaming de áudio para várias redes e dispositivos	Negociação dinâmica de largura de banda	A qualidade musical é aceitável, mas não é excelente	Não



### 3.2- O Uso dos Formatos de Áudio nos Tipos de *Websites*: Empresariais-comerciais, Rádios-WEB, Institucionais e Pessoais

A multiplicidade de usos dos formatos de som na Internet é muito grande. Entre eles podem ser citados os seguintes:

- a) **os fundos musicais e efeitos de som** (como o som do apertar de um botão ou o deslizar de um menu), que podem servir como reforçador de *feedback* no processo interativo do usuário com o *site* na Internet, e também como uso puramente decorativo ou artístico (exemplo: um som de ondas quebrando como fundo sonoro de um *site* de surfe);
- b) **os banco de sons históricos**, onde se tem bancos de sons de documentação histórica disponível *on-line*, com todo o tipo de sons, como sons de acontecimentos históricos, discursos políticos, vozes de grandes cientistas e mais milhares de outros documentos sonoros (este tipo de uso é bem demonstrado no *site* "Virtual Gramophone" – <http://www2.nlc-bnc.ca/gramophone/src/abtdb.htm> – do Canadá que tem um banco de dados multimídia com gravações históricas de discos de 78 rpm do início do século XX, sendo que o visitante pode ouvir estas gravações históricas em RealAudio);
- c) **os acontecimentos jornalísticos**, onde pode-se ter acesso *on-line* de entrevistas e pacotes de notícias dos mais variados tipos;
- d) **os documentos educacionais**, onde aparecem os tutorais, livros e aulas sonoras de múltiplas formas;
- e) **os usos em publicidade**, onde os anunciantes usam todas as técnicas para vender o seu produto com o auxílio do som;
- f) **as rádios-WEB**, onde além das rádios comerciais, tem-se que um usuário comum pode montar a programação de sua própria rádio (na realidade o usuário não cria

aqui uma rádio, mas apenas monta uma *playlist*: lista de músicas a serem tocadas) baseada em fluxos de *streaming* múltiplos que são gerados continuamente em *sites* especializados. Contudo, já é realmente possível comprar pacotes que montam um sistema co-servidor no PC de um usuário caseiro, fazendo com que qualquer um possa montar uma verdadeira rádio-WEB completa em casa (tocando MP3 e outros formatos, além da retransmissão da entrada da placa de áudio com até a própria voz do usuário) como é realizado pelo sistema da empresa *Destiny Media Technologies* ([www.pirateradio.com](http://www.pirateradio.com));

g) **webfonia, audioconferência e videoconferência**, que são tecnologias de comunicação em tempo real.

O poder de informação disponível atualmente na Internet cresceu em grande escala. Uma das formas mais revolucionárias do uso do som na WEB é na forma de livros sonoros (*audio books*), onde encontramos *sites* que comercializam os livros mais variados na forma de som, dentre eles, até mesmo *best-sellers* como o livro de Stephen Hawking "Uma Breve História do Tempo" no formato RealAudio e Windows Media no site [www.audible.com](http://www.audible.com) (este é apenas um pequeno exemplo). Os preços, em geral, destes livros em som são muito baixos (não há custo de papel, impressão e outros mais, comuns ao livro em papel). A aceitação deste tipo de produto parece que tende a aumentar, pois existem muitas pessoas que não gostam de ler ou não têm tempo para tal, mas gostariam de ouvir um livro. Não é preciso imaginar muito o que significa para os deficientes visuais ou outras pessoas com deficiências mais graves, como a tetraplegia, este tipo de tecnologia ganhar força no mercado como produto de grande vendagem e de fácil consecução. Além dos livros sonoros, existem muitos outros produtos falados: discursos históricos de celebridades, aulas gravadas que podem ser acessadas a qualquer momento por um estudante, sendo a parte visual seguida em seu livro didático em casa (podendo também este ser um *e-book* no formato PDF ou DOC).

### 3.2.1- Empresariais-comerciais

Nos *sites* do grupo empresariais-comerciais (os *sites* dos 4 grupos foram obtidos a partir do buscador [www.google.com](http://www.google.com): ver Metodologia) observou-se o Flash como formato mais comum, sendo encontrado em 15,8% dos *sites* vistos. Em 75% dos *sites* não havia som de qualquer formato (Figura-20).

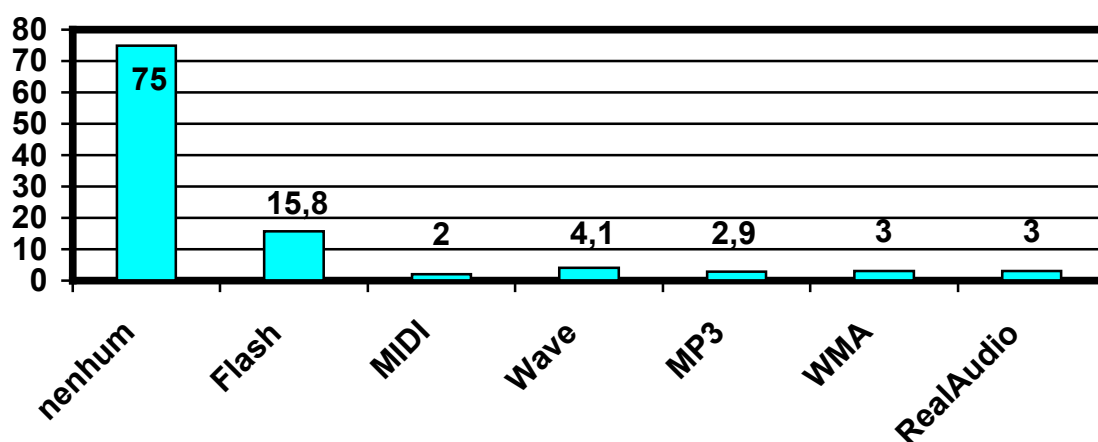


Figura 20 - Ocorrência porcentual (%) dos formatos de áudio nos *sites* amostrados do grupo empresariais-comerciais.

O som do Flash, na verdade, é compactado com um *codec* MP3, sendo ele mais propriamente um formato completo multimídia. A produção de *sites* eficientes com o formato Flash exige técnica, talento e conseqüentemente maiores custos. Esse foi provavelmente um dos motivos de ser o formato preponderante entre as grandes corporações industriais e comerciais.

O formato WAVE foi encontrado em 4,15% dos *sites* (Figura-20) na sua grande maioria como som estrutural na interface (respostas a ações do usuário com o *mouse* e com o teclado: *mouse-down*, *mouse-up* e outros). Estes *sites*, em grande parte, eram de conceitos mais antigos, ainda sem a presença de recursos multimídia mais eficazes (vídeo, animação).

O MP3 foi observado em 2,9% dos *sites* (Figura-20). A grande maioria destes eram de *sites* comerciais de músicas.

Os formatos Windows Media (WMA) e RealAudio apareceram ambos em 3% dos *sites* (Figura-20). Estavam principalmente na forma de música, explicações em áudio sobre produtos, *audio-books* (livros em áudio) e *streaming* de notícias em *sites* jornalísticos.

O MIDI ocorreu em 2% dos *sites* (Figura-20), sendo sempre na forma de fundo musical.

A Figura-21 mostra mais nitidamente as proporções entre os formatos de som encontrados neste grupo, retirando-se o percentual de *sites* que não tinham som e readaptando os *sites* restantes (com som) a uma nova escala, contudo, proporcional.

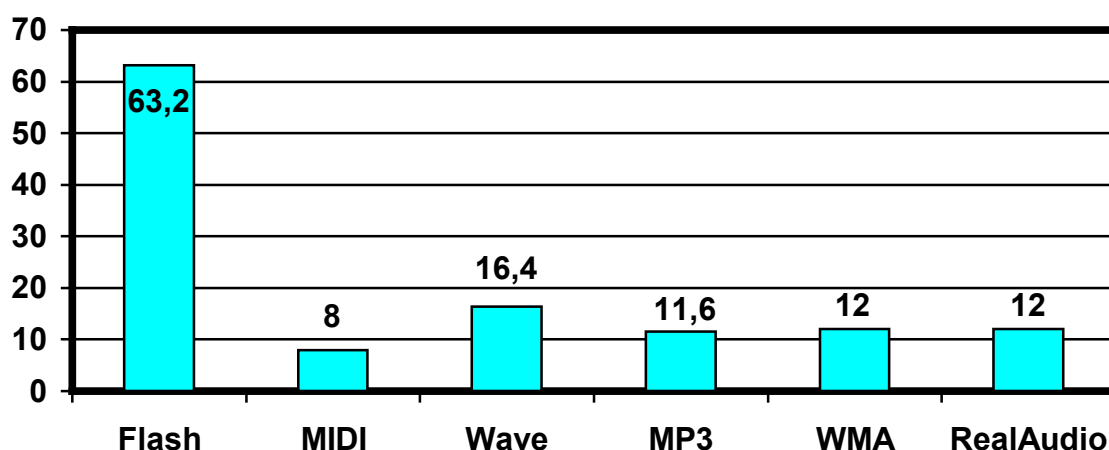


Figura 21 - Presença percentual (%) dos formatos de áudio nos *sites* amostrados do grupo empresariais-comerciais que apresentavam obrigatoriamente algum tipo de áudio. Os *sites* que não possuíam áudio foram desconsiderados nesta representação.

### 3.2.2- Rádios-WEB

Visto que uma rádio-WEB tem obrigatoriamente pelo menos um tipo de formato de som (o *streaming* de transmissão da própria programação), este grupo de *sites* não apresentou no gráfico de ocorrência percentual o item “nenhum” (*sites* com nenhum formato de som) como nos outros 3 grupos (Figura-22).

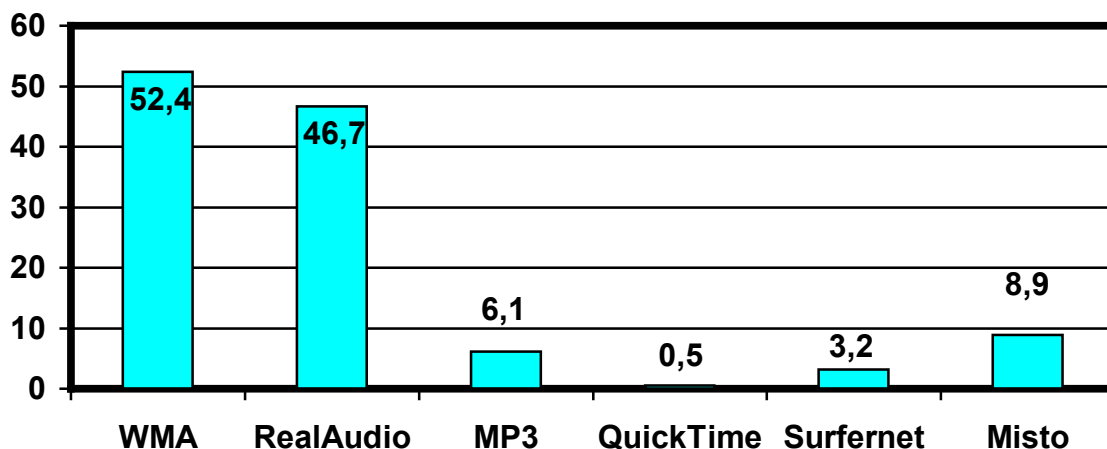


Figura 22 - Ocorrência percentual (%) dos formatos de áudio nos *sites* amostrados do grupo rádios-WEB. Aqui neste grupo apenas os formatos para transmissão da programação (via *streaming*) foram considerados, qualquer outro tipo de áudio estrutural dos *sites* foi desconsiderado.

Existe uma certa tendência em afirmar que o formato da Microsoft (Windows Media Audio: WMA) é o formato que mais cresce em uso na Internet. Uma pesquisa feita em maio de 2002 demonstrava o seguinte padrão de uso na tecnologia *streaming* de transmissão em rádios-WEB pela grande rede:

- 1) WMA----- 52,1%
- 2) MP3----- 34,7%
- 3) RealAudio----- 13,2%

Segundo esta pesquisa, o formato RealAudio vinha em terceiro lugar com apenas 13,25%, o que denotava uma grande queda para o formato que tinha praticamente o monopólio da tecnologia *streaming* no final da década de 1990 (MA-RADIO, 2003).

Entretanto, durante a pesquisa deste trabalho, pôde-se observar um certo equilíbrio entre dois formatos, detectando-se o Windows Media (WMA) em 52,4% dos *sites* analisados, enquanto que o formato RealAudio aparecia em 46,7% dos *sites* (Figura-22).

O formato MP3 vinha depois com 6,1% dos *sites*, enquanto que o QuickTime era o menos comum como rádio-WEB (Figura-22).

Um fato surpreendente foi a aparição na frente do formato QuickTime do sistema de rádios-WEB com tecnologia proprietária “**Surfernet**”, que até pouco tempo atrás não era bem conhecida. Esta tecnologia apareceu em 3,2% das rádios analisadas (Figura-22).

Em 8,9% das rádios encontrou-se a transmissão feita em mais de um formato, contudo, em nenhuma delas havia mais de 2 formatos (Figura-22).

Um fato que não era raro, era encontrar *sítes* de rádios transmitindo em variadas taxas de bits. Por exemplo, em uma certa rádio-WEB foram encontradas as seguintes taxas de bits para o formato WMA: 128, 42 e 28 Kbps. Isto demonstra a necessidade de baixas taxas de bits, visto que a maioria das conexões pelo mundo ainda parecem usar *modems* de até 56 Kbps.

### 3.2.3- Institucionais

Um fato que foi observado nos *sítes* institucionais é o pequeno uso de áudio nos *sítes* ( 89,9% dos *sítes* não continham áudio algum) (Figura-23). Este grupo de sites congrega grandes associações e organizações de todo tipo (clubes, associações, ONG's, cooperativas, instituições de ensino, órgãos do governo).

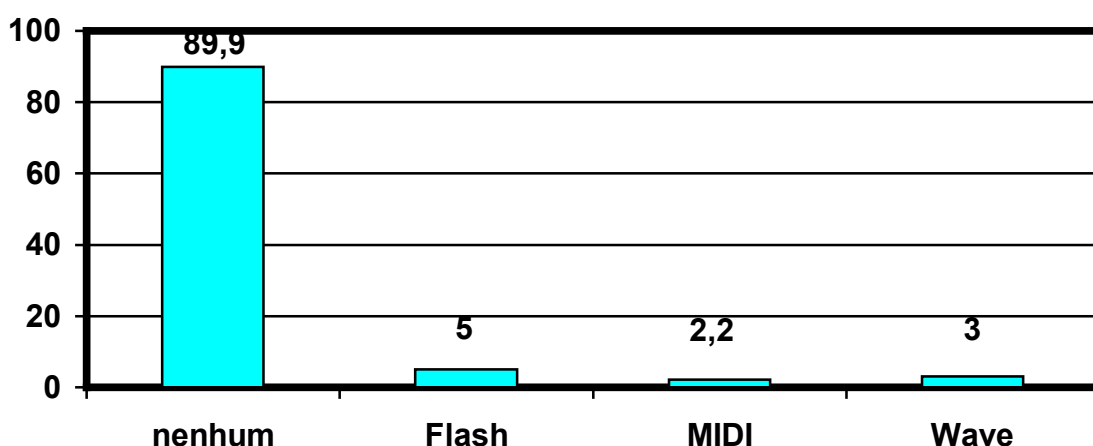


Figura 23 - Ocorrência porcentual (%) dos formatos de áudio nos *sítes* amostrados do grupo institucionais.

O formato Flash foi o mais comum, aparecendo em 5% dos *sites*, enquanto o MIDI apareceu em 2,2% e o formato WAVE em 3% dos *sites*. Aqui o formato MIDI foi utilizado integralmente como fundo musical, enquanto o formato WAVE era utilizado em vários pontos estruturais do *site*: fundo musical, resposta a ações do usuário e conteúdos explicativos.

Na Figura-24 pode-se comparar proporcionalmente a importância dos 3 formatos de áudio encontrados no grupo institucionais: Flash (49,5%), MIDI(21,8%) e Wave(29,7%).

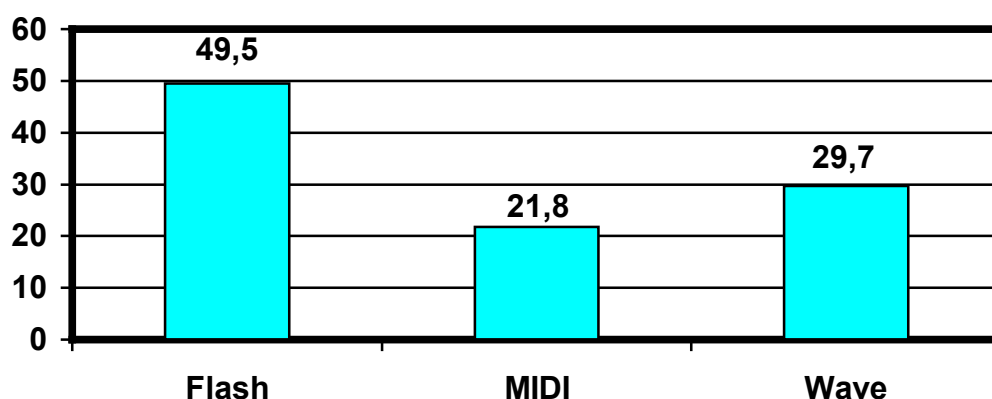


Figura 24 - Presença percentual (%) dos formatos de áudio nos *sites* amostrados do grupo institucionais que apresentavam obrigatoriamente algum tipo de áudio. Os *sites* que não possuíam áudio foram desconsiderados nesta representação.

### 3.2.4- Pessoais.

Muitas pessoas com conhecimento básico na área de informática sabem editar parcialmente a linguagem HTML devido às facilidades dos grandes editores de HTML da atualidade (Homesite, Dreamweaver, Frontpage). Contudo, a inclusão de áudio em um *site* ainda se mostra um tanto quanto técnico para a grande maioria de usuários comuns. Mesmo assim aqui, na amostragem dos *sites* pessoais, encontramos maior presença de *sites* com formatos de áudio (Figura-25) e também de variedade de formatos do que no grupo institucionais.

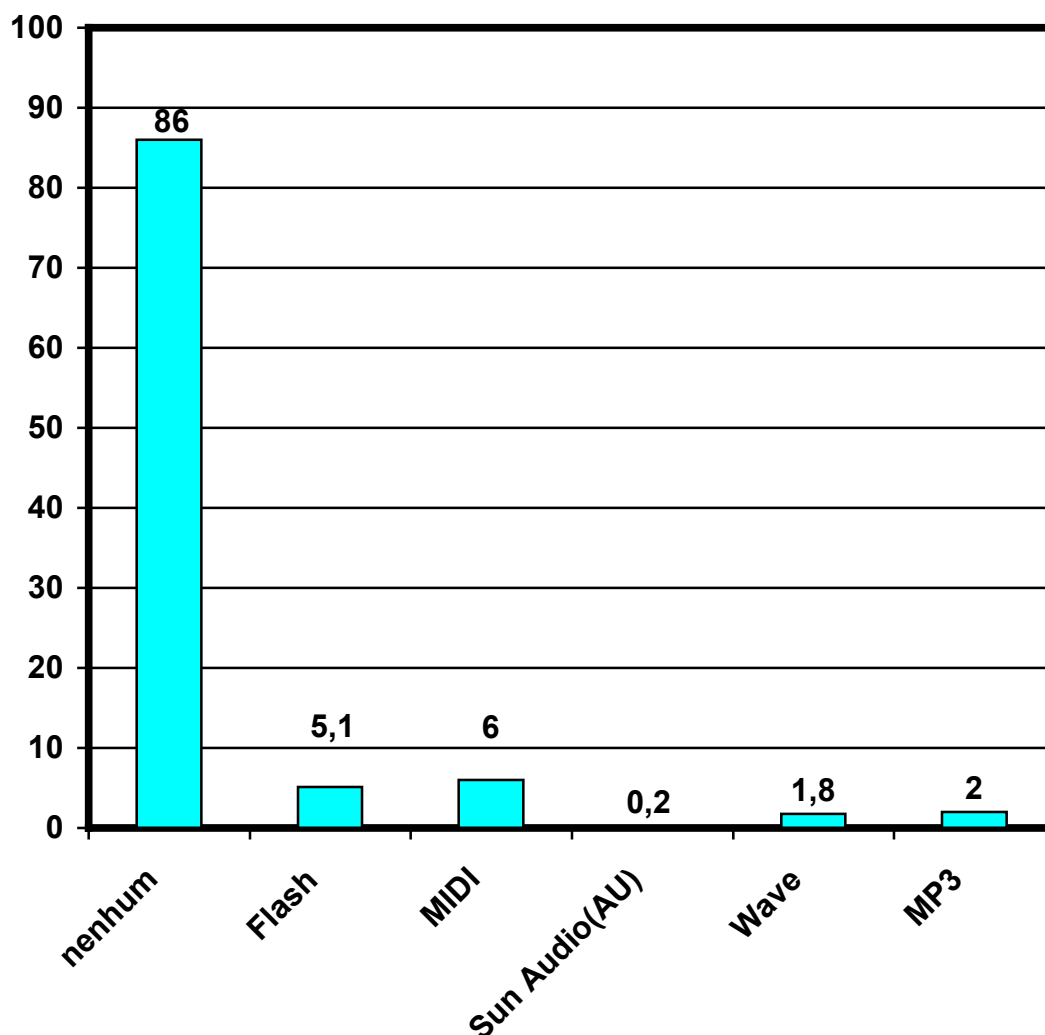


Figura 25 - Ocorrência porcentual (%) dos formatos de áudio nos *sites* amostrados do grupo peessoais.

O que caracteriza basicamente os *sites* pessoais é a grande despreocupação estética, ergonômica, pedagógica e técnica do uso de formatos de som.

O formato mais encontrado nos *sites* pessoais foi o formato MIDI com 6 % (Figura-25), usado aqui como fundo musical e também como forma de distribuição de músicas. Existe pela Internet muitos grupos de apreciadores do formato MIDI. Estas pessoas disponibilizam em seus *sites* pessoais muitas músicas neste formato, colecionando e até trocando arquivos entre si. O MIDI foi um formato que se manteve na WEB graças ao barateamento dos instrumentos MIDI (principalmente teclados) facilmente conectáveis ao computador pessoal, e também principalmente pelo



tamanho pequeno dos arquivos. O aparecimento de *softwares* sofisticados de seqüenciamento MIDI também ajudou a manutenção da popularidade deste formato, pois assim, pessoas comuns chegam a ter verdadeiros estúdios de som virtuais em seus PC's. Outro fator que ajudou o formato MIDI a se manter, foi o aparecimento de placas de som com bons sintetizadores *wave-table*, permitindo assim a execução fiel do som entre diferentes placas de som e plataformas de *hardware*.

Logo após, na pesquisa deste trabalho, apareceu o formato multimídia Flash com 5,1% dos *sites*. Este formato tem se tornado cada vez mais presente em todos os tipos de *sites*, principalmente nos pessoais (Figura-25).

O formato MP3 apareceu em 2% (Figura-25) dos *sites* analisados, sendo esta totalidade composta de música disponibilizada para *download* pelos criadores dos *sites*.

O formato WAVE apareceu com 1,8% de ocorrência, com os usos mais variados nos *sites*: desde fundo musical até elementos da interface (Figura-25).

Aqui neste grupo de *sites* encontramos um formato que é praticamente um resíduo de outras épocas, que é o formato AU da Sun, com 0,2 % (Figura-25). Este era um dos grandes padrões multiplataformas no início da Internet. Era um formato direcionado para o UNIX e seus correlatos.

Na Figura-26 pode-se comparar melhor a ocorrência proporcional entre os 5 formatos de áudio encontrados no grupo pessoais: Flash (36,4%), MIDI(42,8%), AU(1,4%), Wave(12,8%) e o MP3 (14,3%).

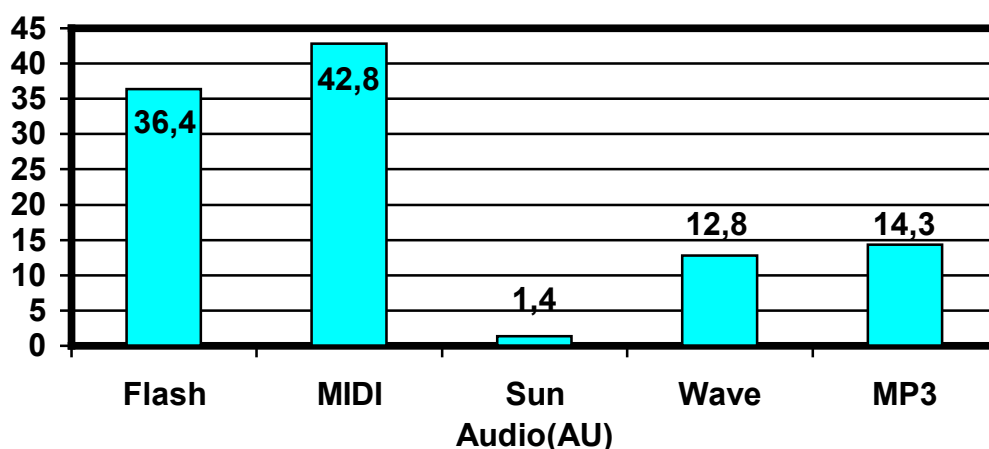


Figura 26 - Presença percentual (%) dos formatos de áudio nos *sites* amostrados do grupo pessoais que apresentavam obrigatoriamente algum tipo de áudio. Os *sites* que não possuíam áudio foram desconsiderados nesta representação.

### 3.3- Disputas, Expectativas e Tendências

Mesmo com grande variedade de formatos de áudio na Internet, apenas alguns deles se sobressaíram e se mantêm firmes na tentativa de se estabelecerem como grandes padrões. Basicamente três estratégias comerciais estão em disputa por estes formatos maiores: a distribuição de arquivos em áudio (*audio-books*, músicas para venda *on-line* e vários outros tipos de pacotes de informação em áudio), o *streaming* de áudio, tanto ao vivo quanto arquivado, e o áudio estrutural que é parte do *design* das interfaces dos *sites* (efeitos de som e fundos musicais). Windows Media Audio (WMA), MP3, Flash, RealAudio e os formatos baseados no *codec* MPEG-2 AAC disputam a maior atenção do mercado. O avanço do formato da Microsoft (WMA) é bem visível. Decorrente disto a empresa RealNetworks (RealAudio) tornou disponível, no início do ano de 2003, o código-fonte de sua tecnologia para a transmissão de áudio e vídeo pela Internet para outras companhias de *software* e *hardware*. Com isso, o formato RealAudio ganha força novamente e tenta recuperar o espaço perdido para o formato WMA.

Apesar de muitas pessoas já afirmarem na qualidade superior do formato WMA (atualmente na versão 9) em relação ao MP3, este último ainda representa a maior fonte de arquivos musicais por toda a rede. Já existe um intrincado sistema de suporte de *hardware* para o MP3: *players* portáteis, *players* para o carro e para casa, DVD's com suporte ao formato e outros dispositivos mais. O suporte de *software* ao MP3 também é bem desenvolvido. Para o formato WMA isto está começando a acontecer, mas de uma maneira mais rápida. Talvez a tarefa da Microsoft não seja tão fácil assim se quiser suplantar o formato MP3.

O formato Flash por ser essencialmente um formato multimídia está fora desta disputa entre WMA, MP3 e RA. A maior vantagem do Flash não é a qualidade (pois não passa de uma forma de MP3), mas sim a sincronização de animações com o som. Com isto o Flash atualmente se estabelece como padrão e assim a empresa Macromedia firma o seu formato, que é posteriormente reconhecido por outras empresas como Apple e Corel que lançam *softwares* similares ao Flash.

A força de um formato também pode ser aumentada ou diminuída graças ao suporte ou não dos produtores de *hardware* e outros dispositivos de som a

determinado formato: placas de som, *players* portáteis, chips dedicados e outros dispositivos usados em casa e no trabalho.

Tal qual o Flash, o formato Beatnik entra na área de interatividade e efeitos de som, contudo o Flash é um formato de animação completo, enquanto que o Beatnik é um de apenas áudio (sem capacidade de animação).

O formato MIDI, apesar de ser um dos mais antigos, ainda mantém uma grande base de usuários pela rede, uma das provas disso é o aparecimento da tecnologia de *MIDI-streaming* “Crescendo”. Grande parte desta base é constituída por adeptos do som MIDI, dado seu pequeno tamanho e facilidade de edição com *softwares* seqüenciadores. Este formato provavelmente vai se manter por vários anos ainda, mesmo fora do âmbito dos músicos profissionais. O formato MIDI não está restrito à Internet, pelo contrário, ele foi desenvolvido para o mundo dos músicos, com as suas composições e outros efeitos.

O formato MOD, que teoricamente seria um formato antigo (plataforma Amiga), ainda se mantém vivo graças a grupos de usuários admiradores. O motivo é que este formato diferente faz com que pessoas comuns sem conhecimento de música tenham a capacidade de compô-las. Este formato deverá perdurar enquanto durar a admiração de seus adeptos, ou até um formato de filosofia semelhante e mais evoluído aparecer no mercado.

O formato Ogg Vorbis, por ser gratuito e de código aberto, com vários desenvolvedores pelo mundo todo, pode futuramente se tornar uma opção à troca de arquivos musicais junto com o MP3 e WMA. Contudo, as últimas novidades desta área demonstram que o grande esforço entre as grandes empresas do ramo musical está sendo direcionado a novas implementações baseadas no MPEG-2 AAC, como é o caso do novo padrão MPEG-4 (que é para se tornar o superpadrão multimídia da Internet). Um dos grandes motivos por esta preferência é a grande qualidade de som conseguida por tal *codec* (MPEG-2 AAC), a sua capacidade *surround* e também pelos seus sistemas de segurança de direitos autorais. De forma geral, é o MPEG-2 AAC que vem se mostrando como o mais promissor entre todos os *codecs*.

Um consórcio de indústrias chamado SDMI ( *Secure Digital Music Initiative*) foi criado pelas gravadoras e outras empresas de interesse da área. Este consórcio pretende pressionar o mercado para o desenvolvimento de formatos que gerenciem direitos autorais e proteção contra cópia. A empresa Microsoft está investindo alto no

seu formato WMA com tais sistemas de proteção. De certa forma, já é meio tarde para esta tentativa ortodoxa. Já existem *terabytes* de arquivos MP3 (que não possuem proteção alguma) dispersos pela Internet e também em mídias gravadas. Contudo, a Microsoft tenta ganhar mercado que é dominado pelo formato MP3 disponibilizando gratuitamente os *encoders* do formato WMA na WEB. Pode-se então criar arquivos musicais WMA sem proteção, tal qual o MP3. Entretanto, é de se esperar que após conquistar o mercado, a Microsoft inicie com o processo de limitação de cópia, devido aos direitos autorais, fato preponderante da queixa das gravadoras no mercado fonográfico. A Microsoft deverá respeitar as reivindicações das gravadoras, pois está fortemente interessada no comércio de música *on-line*.

O maior problema da mudança da filosofia de distribuição dos arquivos de som é que o *hardware* e o *software* atualmente existentes podem não estar preparados para novos formatos de som. Do mesmo modo, os novos equipamentos e *softwares* deverão ser compatíveis com os formatos mais antigos. Tem-se ainda, que grande parte dos sistemas de proteção poderá ser quebrada por programadores talentosos. Alguns especialistas afirmam que as situações podem ser tornar bem mais complicadas com muitas ações judiciais futuras representativas de interesses conflitantes. Como se pode observar, o processo de tornar protegida a distribuição de música pela Internet ainda tem um longo caminho a seguir.

Numa outra perspectiva, agora reportando-se à base tecnológica do som digital, talvez venha-se a assistir ao afastamento do som do computador do padrão do áudio PCM, aproximando-se de novos métodos de gravação. Existem limitações inerentes ao PCM - a qualidade de som pode ser melhorada ao se usar bits extras e frequências de amostragens mais altas, mas estes pequenos ajustes não fornecem uma melhoria linear. Por outras palavras, quanto mais se aumenta a velocidade de bits e de mistura (amostragem), menos se notam os saltos perceptíveis na qualidade de áudio. É muito difícil conseguir ouvir qualquer diferença depois de se passar da resolução de 20 *bits*, pelo que não existe, em concreto, qualquer razão que justifique desenvolver áudio PCM acima de 24 *bits* de resolução e 96KHz de taxa amostral. Talvez o próximo passo possa ser o **Direct Stream Digital**, uma nova forma de gravar som desenvolvida pela Sony para utilização no formato **SuperAudio CD** (criada em conjunto com a **Philips**). O **DSD** continua a misturar som, mas codifica o áudio para o disco de uma forma diferente. Como o sinal de 1 *bit* é misturado a uma

velocidade muito alta, está mais próximo de um fluxo contínuo do que de um sinal PCM convencional. Isto faz teoricamente com que se aproxime mais de um sistema analógico de grande qualidade do que outros sistemas digitais atuais ou antigos. No entanto, os testes de audição entre os consumidores revelaram ser extremamente difícil encontrar a diferença entre reprodução de áudio PCM a 24bits e 96KHz e a reprodução DSD, embora os engenheiros de som tendam a favorecer a tecnologia DSD. Não obstante, o processo ainda está no início e pode muito bem ser que esta se torne a tecnologia preferida para áudio sem perda no futuro (PERSONAL COMPUTER WORLD, 2001a).

De acordo com a empresa de pesquisa pela Internet ARS Inc., por volta de 2004, 25% de toda a mídia será distribuída através de *download* pela Internet, com os arquivos de som e *softwares* liderando este processo (SMARTCOMPUTING, 2001).

O aparecimento dos gravadores de DVD's, o constante avanço dos sistemas de *hardware surround* e a melhoria na velocidade das conexões de Internet levarão a uma intersecção de tecnologias no futuro. A Tecnologia *surround* (multicanais) antes usada somente em cinemas, centro de eventos, aparelhos de som de ponta, DVD's e em placas de som mais modernas para computadores, acabará por se incorporar à Internet, trazendo assim um nível de fidelidade de som jamais conseguido antes pelo usuário da grande rede.

## CAPÍTULO 4- CONCLUSÕES

### 4.1- Conclusões

Devido ao tema de estudo ser muito atual, como são quase todos os assuntos relacionados à Internet, teve-se então que se basear quase que totalmente a pesquisa bibliográfica na própria WEB. Os trabalhos e artigos encontrados sobre formatos de arquivos de som na WEB ou eram extremamente técnicos e específicos ou eram superficiais demais, por isso, foi revista uma grande quantidade de artigos, reportagens e textos técnicos descritivos para arrebanhar material suficiente para o embasamento do tema. Em geral, os artigos somente se referiam aos formatos mais comuns sendo desconsiderados os formatos mais incomuns ou menos usados. Como a quantidade de formatos de arquivos de áudio proprietário no mundo digital é grande, procurou-se identificar e descrever os formatos mais conhecidos e também alguns menos conhecidos. Descrever todos ou sequer a maioria de formatos de áudio digital na WEB seria impraticável para um trabalho de pesquisa com proposta neste nível.

Da pesquisa realizada, foi observado que certos formatos mantiveram-se no mercado até hoje mais por tradição e compatibilidade do que por aplicabilidade tecnológica, como no caso dos formatos WAV (padrão Microsoft), AIFF (padrão Apple) e AU( Unix). Outros formatos antigos mantêm-se até hoje por suas singularidades, como é o caso do formato MIDI, que é uma verdadeira partitura digital que pode ser editada (de sons sintetizados) e do formato MOD, que é uma montagem de amostras reais. Formatos como Beatnik e Crescendo são tecnologias baseadas no formato MIDI, relacionando-se com a interatividade nas páginas WEB no primeiro caso (Beatnik) e com o *streaming* de arquivos MIDI no segundo (Crescendo).

Visto a gama de características específicas de cada formato, os aspectos legais e a competição entre as grandes empresas nesta área da mídia digital, pode-se esperar uma disputa entre alguns poucos formatos em determinadas áreas específicas. Formatos como MP3, MPEG-2 AAC (e seus derivados como Liquid Audio), Flash, MIDI, Windows Media Audio (WMA), Quicktime e RealAudio (RA), parecem já ter se firmado no mercado. Outros formatos, como o QDX, ainda tentam se firmar no meio

comercial. Muitos outros formatos, entre eles o A2B e o VQF, até já saíram do mercado. No caso do VQF poderá até ser continuada a sua implementação, visto tratar-se de tecnologia inovadora em relação a todas as outras. O grande uso comercial destes formatos na Internet mostra a importância e a eficiência de cada um deles em determinada área de atuação da tecnologia de áudio como: *streaming*, *design* e arquivamento de dados sonoros. Contudo, o melhoramento dos formatos existentes e o aparecimento de novas tecnologias, como por exemplo, a *Direct Stream Digital*, está acontecendo com rapidez e a disputa para estabelecer um padrão para determinado fim se torna constante. Em resposta a esta corrida comercial surgiu o formato *Ogg Vorbis*, que é um formato aberto e gratuito, sendo ele mantido e melhorado por vários desenvolvedores de *codecs* pelo mundo todo. Nem sempre na escolha de um formato padrão é eleito o melhor e mais capacitado tecnologicamente, pois existe forte poder de *marketing* de grandes corporações apoiando os seus próprios formatos. Aqui então, o que está em jogo são estratégias muitas vezes mais mercadológicas do que tecnológicas. Outro ponto é que, para que padrões sejam estabelecidos, como os do grupo MPEG, são necessários muitos debates e acordos, com muito tempo e trabalho despendidos. Outra disputa que ocorre na área de áudio na Internet, se situa entre as indústrias fonográficas, músicos e o público consumidor, no direito de escutar, copiar e vender músicas. A disputa pelo comércio *on-line* de músicas e o estabelecimento de formatos para tal fim promete ser vigorosa. Atualmente o formato Liquid Audio é o que se encontra mais em pauta, mas já está sendo seguido pelo Quicktime e pelo WMA.

Na área de rádios-WEB praticamente o formato RealAudio e Windows Media, seguido pelo MP3 estão se estabelecendo como os padrões determinantes. A venda de músicas, as rádios-WEB e WEB-telefonía são provavelmente os maiores filões de exploração econômica com relação ao áudio na Internet (pode-se incluir aqui também a videoconferência com a sua parte de áudio).

O formato que deverá se tornar o superformato multimídia da Internet é o MPEG-4, que tem inovações de várias tecnologias e formatos, entre eles, o do *codec* MPEG-2 AAC, que se mostrou o mais revolucionário tecnologicamente e propenso a futuras implementações. Estabelecer tal superpadrão não é apenas um desafio tecnológico, mas também um desafio mercadológico.

Foi observado que certos formatos foram predominantes nos grupos de *sites* analisados na pesquisa:

**a) EMPRESARIAIS-COMERCIAIS: formato Flash**

**b) RÁDIOS-WEB: Windows Media Audio e RealAudio**

**c) INSTITUCIONAIS: Flash**

**d) PESSOAIS: MIDI e Flash**

Na confecção dos gráficos de barra dos resultados desta pesquisa (seção 3.2) foram desconsiderados os formatos que tiveram ocorrência (nos 4 grupos de *sites*) menor que 0,1%. Estes formatos foram considerados ocasionais, sendo isto feito para facilitar a compreensão estatística e também para melhor visualização nos gráficos.

Pôde-se observar durante a pesquisa neste trabalho que a Internet na atualidade (ano de 2003) parece ser dominada pelos seguintes formatos: Windows Media Audio, RealAudio, Flash, MP3, MIDI e WAVE.

Conforme observado no Quadro-2 (item 3.1.20) pode-se de certa maneira estabelecer o uso mais indicado para cada formato. Assim tem-se que:

- 1) Sonorização interativa de *sites*: Beatnik, Flash/Shockwave.
- 2) Sonorização não-interativa de *sites*: Sun-áudio, WAVE, AIFF, MIDI, Crescendo, MOD.
- 3) Arquivos de músicas: MP3, MPEG-2 AAC, WMA, Ogg Vorbis.
- 4) Venda *on-line* de músicas com proteção de direitos autorais: WMA, Liquid Audio, Quicktime.
- 5) Plataforma CD-ROM: WMA, AIFF, WAVE, MP3, MIDI.
- 6) Arquivos de dados históricos e *audio-books*: MP3, MPEG-2 AAC, WMA, Ogg Vorbis, RealAudio, WAVE, AIFF, Sun-áudio.
- 7) Composição e edição musical: MIDI e MOD.
- 8) Rádios-WEB e *streaming* de notícias: MP3, WMA, RealAudio, QDX, Quicktime, MPEG-2 AAC.



**9) Multiplataforma:** MP3, RealAudio, Sun-áudio, Quicktime, MIDI, AIFF,Flash.

Contudo, toda esta revolução tecnológica com relação ao áudio na Internet só poderá ser posta em prática com a melhoria das linhas de transmissão de dados e conexões de usuários mais rápidas (cabo, ADSL, T1), caso contrário, serviços como WEB-telefonía e rádios-WEB ficarão prejudicados e muitas vezes até inviáveis.

## 4.2- Recomendações para Trabalhos Futuros

- Análise mais pormenorizada dos padrões MPEG, principalmente dos derivados do MPEG-2 AAC, que é o formato que se mostra mais promissor em várias áreas da Internet e até fora dela;
- Análise e teste da relação qualidade-tamanho de formatos menos conhecidos, utilizando-se de equipamentos de qualidade e com pessoas de audição treinada (evitar erros inerentes ao ensaio em si), procurando descobrir algum formato inovador disponível na Internet;
- Análise e teste do comportamento do *streaming* dos formatos mais usados com os variados tipos de conexões domésticas de Internet;
- análise mais pormenorizada dos formatos e *codecs* utilizados nas principais áreas de atuação do áudio na Internet : venda de músicas *on-line*, rádios-WEB e WEB-telefonía;
- discussão sobre se a existência de um formato único para determinado fim seria nocivo ou se aumentaria a eficiência do tipo de serviço.

## REFERÊNCIAS BIBLIOGRÁFICAS

5DDESIGN. **Audio Compression Formats Overview**. *Site comercial*. Disponível em (<http://www.5ddesign.com/html/audioS2.html>). Acesso em 28/01/2001.

ABOUT.COM. **How to Share Files Now -A Guide to the Confusing Alternatives**. *Site educacional*. Disponível em (<http://mp3.about.com/library/weekly/aa071601.htm>). Acesso em 29/11/2002.

ADOBE SYSTEMS. **Beatnik - About RMF**. *Site da empresa Adobe*. Disponível em (<http://www.adobe.com/support/techguides/webpublishing/audio/page3.html>). Acesso em 16/10/2001.

ANDERSSON, H. **Zap's MUSIC-ON-THE-NET Tutorial**. *Site de tecnologia*. Disponível em (<http://www.lysator.liu.se/~zap/tutorial/formats.html>). Acesso em 09/08/2001.

BAGWELL, C. **File Formats**. *Site pessoal*. Disponível em (<http://members.home.com/chris.bagwell/AudioFormats-11.html>). Acesso em 07/08/2001(a).

BAGWELL, C. **Compression schemes**. *Site pessoal*. Disponível em (<http://members.home.com/chris.bagwell/AudioFormats-4.html>). Acesso em 07/11/2001 (b).

BEATNIK. **Beatnik - Basics**. *Site da empresa Beatnik*. Disponível em ([http://www.beatnik.com/software/tutorials/beatnik\\_basics.html](http://www.beatnik.com/software/tutorials/beatnik_basics.html)). Acesso em 20/09/2001(a).

BEATNIK. **Adding sound to the Web**. *Site da empresa Beatnik*. Disponível em (<http://www.beatnik.com/software/tutorials/rmf3.html>). Acesso em 20/09/2001(b).

BEGGS, J.; THEDE, D. **Designing WEB Audio**. CA: O'Reilly, 2001. 382 pp.

BOUVIGNE, G. **VQF**. *Site de tecnologia*. Disponível em ([http://www.mp3-tech.org/top\\_index.html](http://www.mp3-tech.org/top_index.html)). Acesso em 21/08/2001.

CASS, R.; SEYMOUR, M. **Digital audio formats**. *Site Multimedia Teaching Strategies*. Disponível em (<http://web.cs.mun.ca/k12media/resources.formats.sound.html>). Acesso em 07/08/2001.

CENTRAL DA MÚSICA. **Formatos: Real Audio**. *Site musical*. Disponível em ([http://www.centraldamusica.com.br/programas/formatos-real\\_audio.asp](http://www.centraldamusica.com.br/programas/formatos-real_audio.asp)). Acesso em 11/03/2002 (a).

CENTRAL DA MÚSICA. **Formatos: Ogg Vorbis**. *Site musical*. Disponível em (<http://www.centraldamusica.com.br/programas/formatos-ogg.asp>). Acesso em 11/03/2002 (b).

CENTRAL DA MÚSICA. **PlayJ**. *Site musical*. Disponível em (<http://www.centraldamusica.com.br/programas/formatos-playj.asp>). Acesso em 11/03/2002 (c).

CENTRAL DA MÚSICA. **Um em cada cinco CD's vendidos no mundo são piratas, causando prejuízo de U\$ 4 bi às gravadoras**. *Site musical*. Disponível em ([http://www.freemp3.com.br/noticias/noticia\\_dms\\_junho\\_0069.asp](http://www.freemp3.com.br/noticias/noticia_dms_junho_0069.asp)). Acesso em 11/03/2002 (d).

CENTRAL DA MÚSICA. **Gartner Group também concluiu que usuários de MP3 compram mais CD's**. *Site musical*. Disponível em ([http://www.freemp3.com.br/noticias/noticia\\_dms\\_julho\\_0092.asp](http://www.freemp3.com.br/noticias/noticia_dms_julho_0092.asp)). Acesso em 11/03/2002 (e).

CENTRAL DA MÚSICA. **Mp3PRO manterá a liberdade de redistribuição**. *Site musical*. Disponível em ([http://www.centraldamusica.com.br/noticias/noticia\\_abril01\\_0001.asp](http://www.centraldamusica.com.br/noticias/noticia_abril01_0001.asp)). Acesso em 11/03/2002 (f).

CENTRAL DA MÚSICA. **Gravadoras podem sofrer colapso.** *Site musical* . Disponível em ([http://www.freemp3.com.br/noticias/noticia\\_dms\\_setembro\\_0055.asp](http://www.freemp3.com.br/noticias/noticia_dms_setembro_0055.asp)). Acesso em 11/03/2002 (g).

CLIPSTREAM. **Clipstream products page.** *Site comercial.* Disponível em (<http://www.clipstream.com/products/>). Acesso em 05/01/2002.

CODING TECHNOLOGIES. **Coding Technologie's aacPlus called "a clear winner" by world's largest professional broadcast association, the European Broadcasting Union.** *Site de tecnologia* . Disponível em ([http://www.codingtechnologies.de/news/assets/20021009\\_ebutest\\_eng.htm](http://www.codingtechnologies.de/news/assets/20021009_ebutest_eng.htm)). Acesso em 29/11/2002.

DISCTRONICS. **DVD Overview & Features.** *Site de empresa de tecnologia.* Disponível em ([http://www.disctronics.co.uk/technology/dvdintro/dvd\\_intro.htm](http://www.disctronics.co.uk/technology/dvdintro/dvd_intro.htm)). Acesso em 24/06/2003.

DISLEY, A. **Sound Recording Technologies.** *Site pessoal universitário.* Disponível em (<http://www-users.york.ac.uk/~acd114/mustech/record.htm>). Acesso em 06/08/2001.

DOBRIAN, C. **Digital Audio.** *Site pessoal universitário.* Disponível em: (<http://www.arts.uci.edu/dobrian/digitalaudio.htm>). Acesso em 07/11/2001.

DUKE UNIVERSITY. **Audio for Web Pages: Digital Audio File Formats.** *Site educacional* . Disponível em (<http://cit.duke.edu/resource-guides/tutorial-web-multimedia/06-audio-formats.html>). Acesso em 13/08/2001.

DUNN, J. **Understanding digital audio formats.** *Site da empresa Microsoft.* Disponível em (<http://www.microsoft.com/windowsce/uplink/tracks/digaudio.asp?style=5>). Acesso em 07/08/2001.

EARLEVEL. **Digital Audio: Aliasing**. *Site de tecnologia*. Disponível em (<http://www.earlevel.com/Digital%20Audio/Aliasing.html>). Acesso em 06/12/2002.

FEIRA DE CIÊNCIAS. **Acústica (parte 2)**. *Site educacional*. Disponível em ([http://www.feiradeciencias.com.br/sala10/10\\_T02.asp](http://www.feiradeciencias.com.br/sala10/10_T02.asp)). Acesso em 06/12/2002.

FLORIDA CENTER FOR INSTRUCTIONAL TECHNOLOGY. College of Education. University of South Florida. **A Teacher's Guide to Distance Learning**. *Site educacional*. Disponível em (<http://fcit.coedu.usf.edu/DISTANCE/chap4.htm>). Acesso em 12/03/2002.

FRANKE, N. **Multimedia Formats graphics and audio**. *Site de tecnologia*. Disponível em (<http://www.towson.edu/~schmitt/audio/>). Acesso em 09/08/2001.

FRAUNHOFER INSTITUT. **MPEG Audio Layer-3**. *Site de tecnologia*. Disponível em (<http://www.iis.fhg.de/amm/techinf/layer3/index.html>). Acesso em 21/08/2001.

FRIES, B. **The MP3 and Internet**. MD: TeamCom Books, 2000. 302 pp.

GEISLER, G. **A Little Bit About Streaming Audio - How Does It Work?** *Site pessoal universitário*. Disponível em (<http://www.ils.unc.edu/~geisg/archive/310/how.htm>). Acesso em 24/01/2002.

GERTH, B.J. **Audio on the Web - Adding Sound to your Web Page**. *Site pessoal universitário*. Disponível em (<http://web.uvic.ca/~bgerth/workshops/audioonweb.htm>). Acesso em 21/11/2001.

HARARD, F. ; PAULISSEN, D. **Multimedia Mania**. Abacus, 1992. 513 pp.

HASS, J. **Principles of Digital Audio**. *Site educacional*. Disponível em ([http://www.indiana.edu/~emusic/digital\\_audio.html](http://www.indiana.edu/~emusic/digital_audio.html)). Acesso em 03/04/2002.

HELENIO, L. **Tutorial - Midi Karaoke**. *Site de tecnologia*. Disponível em (<http://www.audiodigital.com.br/audiodigital/news/Kar/Karaoke.html>). Acesso em 27/09/2001.

HENRY, J. **Sound Formats**. *Site educacional*. Disponível em ([www.henryjim.com/audio/sound\\_formats.html](http://www.henryjim.com/audio/sound_formats.html)). Acesso em 21/08/2001.

LIEB, T. **Sound on the Web**. *Site pessoal universitário*. Disponível em (<http://www.towson.edu/~lieb/editing/sound.html>). Acesso em 07/08/2001.

LINDSTROM, R.L. **Apresentações em Multimídia**. Makron Books, 1996. 485 pp.

LITEXMEDIA. **Description of most popular audio formats**. *Site comercial*. Disponível em ([http://www.litexmedia.com/article/audio\\_formats.html](http://www.litexmedia.com/article/audio_formats.html)). Acesso em 05/01/2003.

MA-RADIO. **Webradios**. *Site de tecnologia*. Disponível em (<http://www.ma-radio.com/webradios.htm>). Acesso em 20/01/2003.

MENLO PARK EDISON MUSEUM. **Before The Phonograph**. *Site educacional*. Disponível em (<http://www.edisonnj.org/menlopark/birthplace/beforethephonograph.asp>) . Acesso em 28/03/2002.

MULLIN, R. **Hearing**. *Site Discovery.com*. Disponível em: (<http://health.discovery.com/convergence/ultrapregnancy/senses/hearing.html>). Acesso em 07/07/2002.

NETPHONES. **Internet Phones**. *Site de tecnologia*. Disponível em (<http://home.rochester.rr.com/netphones/buddyphone.htm>) . Acesso em 02/04/2002.

NILSSON, M. **The short history of tagging**. *Site ID3V2 de tecnologia*. Disponível em (<http://www.id3.org/history.html>). Acesso em 21/08/2001.

NUNTIUS. **Multimedia - Streaming Audio -Codecs**. *Site de tecnologia* . Disponível em (<http://www.nuntius.com/solutions31.html>). Acesso em 28/01/2002.

PATTERSON, J.; MELCHER, R. **Audio on the WEB**. Berkeley, CA: Peachpit Press, 1998. 204 pp.

PERSONAL COMPUTER WORLD. **Do analógico ao digital**. *Site de tecnologia*. Disponível em (<http://www.fbnet.pt/pcw/0301/a03-02-00.shtml>). Acesso em 30/10/2001(a).

PERSONAL COMPUTER WORLD. **Do analógico ao digital**. *Site de tecnologia*. Disponível em (<http://www.fbnet.pt/pcw/0301/a03-00-00.shtml>). Acesso em 30/10/2001(b).

PIERCE, C.A. **How It All Began...the History of Napster**. *Site pessoal universitário*. Disponível em (<http://holly.colostate.edu/~cass859/allbegan.html>). Acesso em 05/01/2003.

PISTORI, J. **Visão Geral de compressão de áudio e vídeo para o uso em videoconferências** . *Site pessoal universitário*. Disponível em (<http://www.inf.ufsc.br/~jefepist/ine6403/jeferson>). Acesso em 03/07/2003.

PROJECTCOOL. **The .ea format**. *Site de tecnologia*. Disponível em (<http://www.projectcool.com/developer/audioz/AZquick/formats/ea.html>). Acesso em 19/12/2001.

QDESIGN. **QDX by QDesign**. *Site da empresa QDesign*. Disponível em ([http://www.qdesign.com/products/about\\_products/qdx.htm](http://www.qdesign.com/products/about_products/qdx.htm)). Acesso em 21/12/2001.

RADIFIED. **Monkey's Audio Lossless Compressor**. *Site de tecnologia*. Disponível em (<http://mp3.radified.com/lossless.htm>). Acesso em 02/01/2003.

REVISTA INFO EXAME. Ed. Abril ,ano 14, nº 163, outubro/1999. 178 pp.



RUBINSTEIN, G. **Audio Recording: History and Development**. *Site Jones Telecommunications & Multimedia Encyclopedia* . Disponível em (<http://www.digitalcentury.com/encyclo/update/audiohd.html>). Acesso em 06/08/2001.

SCHERER, W.W. **MID versus WAV& RA, Which sound files to use in Web pages**. *Site comercial musical*. Disponível em (<http://www.midiloops.com/midvswav.htm>). Acesso em 09/08/2001.

SCHOENHERR, S. **Recording Technology History**. *Site educacional*. Disponível em (<http://history.acusd.edu/gen/recording/notes.html>). Acesso em 06/08/2001.

SERIOUS CYBERNETICS. **A comparison of Internet audio compression formats**. *Site de tecnologia*. Disponível em (<http://www.sericyb.com.au/audio.html>). Acesso em 09/05/2003.

SMARTCOMPUTING. **What Is Digital Audio?** *Site de tecnologia*. Disponível em (<http://www.smartcomputing.com/editorial/article.asp?article=articles%2Farchive%2Fg0904%2F38g04%2F38g04%2Easp>). Acesso em 07/11/2001.

SIMPSON, R. **Cutting Edge WEB Audio**. NJ: Prentice Hall, 1998. 536 pp.

TERRAVISTA. **Outros codecs de Audio**. *Site educacional*. Disponível em (<http://www.terravista.pt/PortoSanto/8716/outros/outros.htm>). Acesso em 11/03/2002.

THB.MUSIC SERVICES. **Sound, Video and Animation FAQ**. *Site de tecnologia*. Disponível em (<http://www.audiohost.com/sound.html>). Acesso em 27/08/2001.

THEODRIC TECHNOLOGIES LLC. **Radio-Locator**. *Site de procura de rádios-WEB*. Disponível em (<http://www.radio-locator.com/>) . Acesso em 25/11/2002.

THE PHYSICS CLASSROOM. **Sound Properties and Their Perception**. *Site* educacional. Disponível em: (<http://www.physicsclassroom.com/Class/sound/U11L2c.html>). Acesso em 28/03/2002(a).

THE PHYSICS CLASSROOM. **Sound Properties and Their Perception**. *Site* educacional. Disponível em: (<http://www.physicsclassroom.com/Class/sound/U11L2a.html>). Acesso em 28/03/2002(b).

THE PHYSICS CLASSROOM. **Sound Properties and Their Perception**. *Site* educacional. Disponível em: (<http://www.physicsclassroom.com/Class/sound/U11L2b.html>). Acesso em 28/03/2002(c).

THE SONIC SPOT. **File Format List**. *Site* de tecnologia. Disponível em (<http://www.sonicspot.com/guide/fileformatlist.html>). Acesso em 27/11/2002.

THOMASON, R. **3 Music Formats Challenge the Ubiquitous MP3**. *Site* de tecnologia. Disponível em (<http://www.washtech.com/news/media/11136-1.html>). Acesso em 21/12/2001.

VILELA, A.L.M. **O mecanismo da audição**. *Site* “Anatomia e Fisiologia Humanas”. Disponível em (<http://www.afh.bio.br/basicos/Sentidos4.htm>). Acesso em 09/05/2003.

VQF.COM. **What is VQF ?** *Site* de tecnologia. Disponível em (<http://www.vqf.com/faq.php3>). Acesso em 13/08/2001.

WHITTLE, R. **Lossless audio compression**. *Site* de tecnologia. Disponível em (<http://www.firstpr.com.au/audiocomp/lossless/>). Acesso em 02/01/2003.

WIKIPEDIA. **MP3**. *Site* educacional. Disponível em (<http://www.wikipedia.org/wiki/Mp3>). Acesso em 09/05/2003.

WILLEBEEK-LEMAIR, M.H.; KUMAR, K.G.; SNIBLE, E.C. **Bamba: Audio and video streaming over the Internet**. Site da empresa IBM. Disponível em (<http://www.research.ibm.com/journal/rd/422/willebeek.html>) . Acesso em 07/12/2001.

WILLRICH, R. **Apostila de Aula: Sistemas de Multimídia Distribuídos**. UFSC- Departamento de Informática e Estatística - Curso de Pós-graduação em Ciências da Computação, junho/2000. 247 pp.

ZHANG, A. **Multimedia File Formats on the Internet: Sound and Music**. Site de tecnologia. Disponível em (<http://www.proxy.rhbnc.ac.uk/Music/Prague/soundfls.html>). Acesso em 09/08/2001.

## GLOSSÁRIO

**ADSL** – Abreviatura de *Asymmetric Digital Subscriber Line*. É um tipo de conexão com a Internet que usa linha telefônica convencional para a comunicação rápida de dados. Diferentemente de um modem comum de 56 Kbps, a tecnologia ADSL pode carregar dados na rede com velocidade de até 640 Kbps e descarregar com velocidades maiores que 6 Mbps. Uma linha ADSL usa a parte da largura de banda da linha telefônica não utilizada pela voz, permitindo transmissão simultânea de voz e dados. Como diz a sua sigla, ela é assimétrica, pois descarrega dados muito mais rápido do que carrega, visto que esta, em geral, é a vontade de um usuário comum.

**Algoritmo** – Conjunto de instruções, como uma receita, que realiza funções ou resolve problemas matemáticos.

**Aliasing** – *Aliasing* é o que acontece quando um sinal sonoro é amostrado (digitalizado) muito lentamente. Especificamente, se um sinal contém frequências acima de metade da taxa de amostragem, a versão amostrada do sinal conterá frequências que não estão no sinal original (*aliases*). Este limite de metade da taxa de amostragem é conhecido como o limite de *Nyquist*.

**Amiga** – Linha de computadores criada pela empresa Commodore (originada nos EUA e crescida no Canadá). Usava chips Motorola. Nos anos 80 ganhou uma legião de aficionados pela sua grande capacidade gráfica e de som para a época. A Commodore, que chegou a ser uma das maiores empresas de microcomputadores do mundo, foi a falência e vendida em 1995.

**Audio-books** – Livros literários gravados na forma de arquivos de som que são comercializados pela Internet. São encontrados principalmente no formato RealAudio.

**Áudio estrutural** – É o áudio (arquivos de som) que faz parte da interface da página WEB, como fundo musical ou como resultado da atividade interativa com os botões e outros elementos da interface.

**Broadcasting** – Em sentido mais geral, "broadcasting" se refere a atividade de rádio e teledifusão. Na parte relativa à Internet, este termo descreve a distribuição simultânea, em tempo real, de qualquer mídia, de uma só fonte para muitos destinatários.

**Browser** – O mesmo que navegador, os dois mais utilizados são: Internet Explorer e Netscape; contudo o uso do Internet Explorer da empresa Microsoft tem se tornado cada vez mais preponderante.

**CCITT** – Sigla para *Comite Consultatif Internationale de Telegraphie et Telephonie*. É um comitê baseado em Genebra, Suíça, que determina padrões em telecomunicações, incluindo *codecs* de áudio e os famosos padrões V. para velocidade e compressão de modems (V.34 e outros). Embora esta organização tenha trocado seu nome para **ITU-T** (*International Telecommunications Union-Telecommunication*), o antigo nome francês sobrevive.

**Codec** – Abreviatura para "*compression/decompression algorithm*". Como diz o nome, os *codecs* são utilizados para codificar e decodificar (ou comprimir e descomprimir) vários tipos de dados – particularmente aqueles que de outra maneira usariam enorme espaço em disco, tais como arquivos de som e vídeo. *Codecs* podem ser usados com tecnologia *streaming* (som e vídeo ao vivo ou não) ou com arquivos gravados. Eles podem ser implementados em *hardware* ou em *software*. Entre os *codecs* mais conhecidos podemos citar: MP3, WMA e RealAudio.

**Compactação com perda** – Tipo de compactação em que o som compactado não contém os mesmos dados do som original. Certos dados foram perdidos. O nível de perda depende do tipo de *codec* e da sua configuração. A vantagem deste tipo de compactação é de que o tamanho do arquivo resultante é, em geral, bem menor do que quando utilizada a compactação sem perda.

**Compactação sem perda** – Tipo de compactação em que o som compactado contém os mesmos dados do som original.

**Criptografia** – É a ciência da segurança da informação. Trata das técnicas para esconder e proteger informações armazenadas ou em trânsito, como cifragem de texto ou mistura de palavras com imagens.

**Decibéis (decibel)** – Uma unidade de medida de altura do som (pressão do som) e alterações ou diferenças no nível do sinal sonoro. O decibel é uma função matemática logarítmica (log) que reduz valores numéricos grandes em valores menores e mais compreensíveis. O nome da unidade é em homenagem ao inventor Alexander Graham **Bell**.

**Dolby** – A Dolby Laboratories, empresa fundada em 1965 pelo norte-americano Ray Dolby, acabou por se tornar uma das empresas líderes em tecnologia de som, tanto para aparelhos de uso caseiro (home theater, DVD), tanto quanto no ramo comercial (filmes e cinemas).

**DSD** – Sigla para “*direct stream digital*”, um formato inovador de codificação de áudio digital proposto pelas empresas Sony e Philips e usado no novo formato de mídia Super Audio CD (SACD). O DSD usa amostras de 1 bit feitas com uma frequência de 2.88 MHz, sendo que com esta altíssima taxa de amostragem, teoricamente, devemos ter uma maior aproximação com o sinal analógico original de fluxo contínuo. Esta tecnologia comprime os dados pela metade.

**Download** – Termo em inglês, mais usado que na forma portuguesa, significa baixar um arquivo da Internet.

**DTS** – Tecnologia de som desenvolvida pela empresa *Digital Theater Systems Inc.* para sistemas de *home theater*, DVD e filmes de cinema. Empresa concorrente da Dolby Labs.

**E-book** – Livro que está totalmente na forma digital e que é vendido comumente pela Internet. Estes livros podem ter vários formatos proprietários, contudo, os padrões mais usados são o “.DOC” (Word da Microsoft) e “.PDF” (da empresa Adobe).

**Encoder** – Parte de um *codec* que faz a codificação (compressão).

**Encriptação** – veja **Criptografia**.

**Erro de quantificação** – Distorção que resulta do processo de quantificação durante a codificação digital do som. A quantificação é um processo em que uma faixa de valores contínuos de um sinal de som análogo é amostrada e dividida em subfaixas não-sobrepostas, e um valor discreto e único é atribuído a cada subfaixa. Este processo de quantificação é utilizado no áudio PCM (*pulse-code modulation*). Se a amostra do sinal quantificado cai dentro de uma dada subfaixa, esta amostra recebe o valor discreto correspondente àquela subfaixa para propósitos de modulação e transmissão.

**Firewalls** – Dispositivo, em geral de *software*, direcionado a proteger servidores de rede, de ações nocivas externas; como por exemplo, a ação de *hackers*.

**Freeware** – *Freeware* é um estilo de *software* que se pode descarregar, passar adiante e distribuir sem necessidade de comprá-lo. Contudo ele ainda tem direitos autorais, de modo que você não pode descompilá-lo ou vendê-lo como se fosse seu (como é permitido no caso de um programa de domínio público).

**Frequência ou taxa de amostragem** – É o número de amostras que são feitas por segundo para representar o som analógico fonte de uma forma digital. As amostras podem ter variadas resoluções de bits como 4, 8, 16 ou 24 bits. Quanto maior a taxa de amostragem mais exata será a representação digital do som.

**Glitches** – Tipo de distorção de rápida duração no som digital.

**Hardware proprietário** – *Hardware* com tecnologia de propriedade de uma certa empresa.

**Homepage** – Página inicial de um *site* na Internet.

**ID3 Tag** – É um tipo de metadata do arquivo MP3 que permite informar elementos como : detalhes do artista e do álbum, letra da música, informações da trilha musical, gênero e outras características mais. Existem *softwares* que editam tais campos. Com os campos *ID3 tags* as músicas em MP3 podem ser melhor organizadas e identificadas pelo usuário.

**Instituto Fraunhofer** – Famoso instituto alemão de pesquisas tecnológicas, uma das entidades responsáveis pelo desenvolvimento do *codec* MP3 (MPEG-1 *layer*-III)

**IP** – O *Internet Protocol* (IP) é o método ou protocolo através do qual os dados são enviados de um computador ao outro na Internet. Cada computador tem pelo menos um endereço de IP que o diferencia de todos os outros computadores da rede.

**IRC** – O *Internet Relay Chat* (IRC) é um sistema para *chat* (conversa entre vários usuários ao mesmo tempo, através do modo texto) que envolve uma série de regras, convenções e *software* cliente/servidor especializado. Esta tecnologia é muito popular entre os jovens.

**ISDN (Integrated Services Digital Network)** – Traduzido por RDSI (Rede Digital de Serviços Integrados) é a digitalização da rede telefônica para tráfego simultâneo de voz, dados, imagens, aplicações e serviços multimídia. O ISDN foi concebido para substituir a rede telefônica convencional (analógica) por uma rede digital. Atualmente também é utilizado como uma solução para acesso à Internet a velocidades mais altas, e a um custo acessível.

**ISO** – Organização não-governamental fundada em 1947 e localizada em Genebra (Suíça), sendo uma federação mundial de entidades de padronização de cerca de



147 países. Dentre estas entidades temos a norte-americana ANSI (*American National Standards Institute*) e a brasileira ABNT (Associação Brasileira de Normas Técnicas).

**ITU** – veja **CCITT**.

**Java** – Linguagem de programação desenvolvida pela empresa Sun Microsystems para adicionar animações e outras ações a *websites*. É uma linguagem orientada a objetos, multiplataforma e direcionada a Internet.

**Jingles** – Pequenas composições musicais, em geral de fácil memorização, que são cantadas com objetivos publicitários.

**Kbps** – Quilobits por segundo (milhares de bits por segundo). É uma medida da largura de banda (quantidade de dados que podem fluir por unidade de tempo) em um meio de transmissão de dados.

**KHz** – Equivale a mil Hertz, que é a unidade de frequência em uma onda sonora. Um Hertz é um ciclo ondulatório por segundo.

**Largura de banda** – Largura de banda é a quantidade de dados que podem fluir por unidade de tempo em um determinado meio de transmissão.

**LP** – Abreviatura para *Long Play*; como eram chamados os padrões dos discos musicais de vinil de 12 polegadas.

**Linguagem "C"** – É uma linguagem estruturada que tem sido grandemente utilizada no desenvolvimento de *softwares* e sistemas operacionais. Muitas versões de sistemas operacionais baseados no UNIX foram escritos em C.

**Macromedia Director** – Famoso programa de geração de conteúdo multimídia. Possui uma linguagem interna chamada Lingo. O formato Shockwave nada mais é do que o formato Director quando compactado para a Internet.

**Marca d'água (*watermarking*)** – Como marca-d'água em um papel comum, a marca d'água digital é um marcador de identificação embutido que não pode ser removido de um documento digital. Ela contém dados escondidos de identificação que são usados para monitorar o uso de materiais com direitos autorais.

**Metadata** – Dados que dão informações auxiliares sobre uma fonte. No caso do som digital, são campos auxiliares como o caso do ID3 tag do arquivo MP3.

**Mbone** – Abreviatura de (*multicast backbone on the Internet*), sendo uma extensão da Internet desenhada para suportar *multicasting* em IP (transmissão de pacotes de dados para múltiplos endereços). Muitas destas transmissões são de *streaming* de áudio e vídeo, exatamente como ocorre com as rádios e TV's, sendo estas transmissões enviadas para muitas pessoas de uma vez só.

**MIDI** – Abreviatura para *musical instrument digital interface*. Padrão de dispositivo que permite a conexão de computadores, instrumentos musicais e sintetizadores para compor e tocar som digital. O termo é utilizado para descrever o padrão em si, o *hardware* que suporta o padrão e os arquivos que armazenam dados compatíveis com o padrão. Os arquivos MIDI são como partituras digitais, eles contêm as instruções para as notas musicais, o tempo e o tipo de instrumento.

**MOD** – Contração de *Module*. Este formato de arquivo combina aspectos de som MIDI e WAVE (som amostrado). Ele contém uma espécie de partitura digital como o MIDI, contudo, em adição, existem várias amostras de som (tipo WAVE) com instrumentos tocando várias notas. Desde que um arquivo MOD contém todas as amostras necessárias para tocar as músicas, este tipo de arquivo musical deve tocar da mesma forma independentemente da placa de som do computador, fato que não ocorre com o MIDI.

**MP3** – Sigla para MPEG *layer-III*. *Codec* de áudio que iniciou a grande revolução do áudio na Internet. É um dos grandes formatos, junto com o WMA, Real Audio, Flash e

MPEG-2 AAC. Apresenta grande suporte de *software* e *hardware*. Em torno dele está uma grande polêmica que envolve o assunto pirataria de músicas pela grande rede.

**MPEG** – Abreviatura de *Moving Picture Experts Group*. Um grupo de trabalho da ISO/IEC encarregado do desenvolvimento de padrões para representação codificada de vídeo e áudio digital. Estabelecido em 1988, o grupo já produziu *codecs* famosos como o MPEG-1(usado no Video CD e no som MP3), o MPEG-2 (TV digital e o DVD) e o MPEG-4 (superpadrão multimídia para a Internet, em finalização).

**ms** – Unidade de tempo: milissegundo (milésimo de segundo).

**NEXT** – A NeXT foi uma empresa de computadores fundada por Steven Jobs, um dos fundadores da Apple Computer, e também o nome do avançado computador pessoal para trabalho (*workstation*) que esta companhia desenvolveu em 1988. Em 1993 a produção deste tipo de *workstation* foi encerrada.

**Multiplataforma** – A habilidade de um produto (*hardware* ou *software*) de suportar uma variedade de plataformas de computadores (e.g. IBM, UNIX, Macintosh, Windows, Amiga); também chamado de plataforma-cruzada.

**Newsgroup** – São grupos de discussão entre os usuários da Internet, organizados por temas, onde se pode ler e escrever (publicamente) sobre o tema indicado pelo nome do grupo de discussão.

**PCM (Pulse Code Modulation)** – Método padrão de digitalização de som sem perda. É o *codec* nativo dos formatos WAVE e AIFF.

**PDA** – Abreviatura para *Personal Digital Assistant*. Termo criado pela Apple quando lançou o dispositivo de mão Newton no início dos anos 90. No início, os primeiros PDA's eram apenas agendas eletrônicas sofisticadas, contudo, hoje em dia, são praticamente computadores completos.

**Peer-to-Peer** – Rede sem um computador (servidor) centralizador. Os usuários da rede compartilham arquivos diretamente entre si.

**Player** – *Software* utilizado para rodar determinado formato de mídia (áudio, vídeo, 3d, animação, texto).

**Plug-in** – Um *plug-in* aumenta as capacidades de um *browser*, tais como o Netscape ou o Internet Explorer, permitindo que estes *browsers* rodem determinados arquivos ou realizem determinadas ações. O Internet Explorer além de ser compatível com muitos *plug-ins* (padrão do Netscape), usa uma tecnologia exclusiva chamada ActiveX.

**Pops** – Tipo de distorção de rápida duração no som digital, resultado de uma codificação mal feita. Distorção muito comum no som MP3.

**Profundidade de bits** – O número de bits utilizado para definir uma amostra. Também é conhecido como resolução da amostra. Quanto maior a quantidade de bits melhor a qualidade de som, contudo, estudos comprovam que acima de 24 bits, em geral, seria desnecessário.

**Programas cliente** – Programas usados pelo lado cliente (ou usuário) de um sistema cliente/servidor. Um típico exemplo seria um *software* de FTP para baixar arquivos da Internet.

**Protocolo** – Devido ao fato de variados tipos de computadores e sistemas operacionais se conectarem entre eles por variados meios dentro da Internet, eles têm de seguir regras de comunicação que torne possível a troca de dados entre eles. Estas regras são denominadas de protocolos. Entre eles podem ser citados : PPP, TCP/IP, SLIP e FTP.

**Psicoacústica** – Ciência que estuda como os sons são interpretados pelo cérebro humano.

**RA** – Contração de RealAudio, formato de som digital.

**Rádio-WEB** – Rádio que tem sua programação divulgada pela Internet através da tecnologia *streaming*. Uma rádio-WEB pode ser ao vivo ou gravada.

**Rebuffering** – O *buffer* é uma memória intermediária auxiliar que facilita as tarefas, pois liberta o processador. *Rebuffering* é o ato de colocar novamente os dados no *Buffer*.

**Real SMIL** – Tecnologia desenvolvida pela empresa Real Networks que capacita a tecnologia Real Media com a linguagem SMIL (ver **SMIL** no Glossário).

**RPM** – Rotações por minuto, medida utilizada para medir o número de rotações que um dispositivo tocador de uma mídia em disco (digital ou analógica) consegue atingir.

**Ruído** – Termo que descreve os sinais que os seres humanos consideram como portadores de pouca informação útil ou aqueles os quais são considerados desagradáveis.

**Sample (amostra)** – O termo refere-se ao menor pedaço de dados usado para representar um sinal de áudio analógico de uma forma digital. Também pode se referir a pequenos trechos de composições sonoras variadas.

**Sampler** – Dispositivo de *hardware* ou *software* que utiliza *samples* (amostras) para gerar uma saída de áudio. Um *sampler* freqüentemente é usado para juntar amostras (*samples*) e criar reproduções realísticas de sons e instrumentos musicais, gerando literalmente novas composições musicais.

**Seqüenciadores** – Um dispositivo de *hardware* ou *software* que é usado para gravar ou ordenar comandos MIDI e outros eventos de áudio em padrões e composições musicais. O *software* Cakewalk é um típico exemplo.

**Servidor de áudio** – Um servidor de áudio é um computador configurado com um *software* servidor direcionado a distribuir arquivos de áudio pela Internet (o termo servidor é também utilizado para descrever o *software* e o sistema operacional desenhados para rodar o *hardware* de um servidor). Como exemplo típico temos um servidor RealAudio.

**Sinais surround** – Som multicanal que deve ser ouvido em sistema de caixas próprias para *surround*. O termo vem de “circundar” ou “ficar ao redor”. A audição humana no mundo real tem a capacidade de identificar de uma certa forma a direção e velocidade de uma fonte de som; quase como se pudéssemos “enxergar” com os ouvidos. A tecnologia *surround* tenta propiciar esta sensação ao ouvinte, fato que no modo estéreo não é conseguido. A tecnologia *surround* é a empregada nos sistemas de *Home Theater*.

**Sintetizador** – Dispositivo que eletronicamente recria sons de outros instrumentos, como pianos, flautas, órgãos e muitos outros. Os sintetizadores mais comuns usam um teclado para emitir sons, contudo os de guitarra estão se tornando cada vez mais populares.

**SMIL** – É uma linguagem de programação do estilo da HTML proposta pelo *World Wide Web Consortium* (WC3) que sincroniza vários tipos de mídias como vídeo, som, texto e imagens.

**SQNR** – Abreviatura para *Signal-to-Quantization Noise Ratio*. É uma medida da qualidade com a qual os sinais análogos de som são convertidos em sinais digitais (quantificação). É expresso em decibéis (dB).

**Streaming** – Tecnologia que permite que uma mídia seja vista pela internet por um computador que ainda não a baixou por completo. À medida que trechos da mídia vão chegando ao computador destino, estes dados vão sendo armazenados em um *buffer* (uma espécie de pré-memória organizadora), para daí serem rodados. Durante

a execução de uma mídia, quando o *buffer* chega a se esvaziar, sem que não haja mais dados para alimentá-lo, então ocorrem saltos e pausas na transmissão.

**T1** – É um termo criado pela empresa AT&T para um tipo de conexão com a Internet que transfere dados digitais a velocidade de 1,544 Mbps.

**T3** – Tipo de conexão digital com a Internet com quase 30 vezes a capacidade de transmissão de dados de uma linha T1 (44,736 Mbps).

**Terabyte (TB)** – Medida de memória equivalente a 1 bilhão de bytes ou mil gigabytes.

**THX** – Abreviatura para *Tomlinson Holman Experiment*. É um conjunto de padrões técnicos desenvolvidos pela Lucasfilm (empresa do diretor George Lucas) em 1982, para assegurar que o público veja e ouça os filmes nos cinemas nos níveis de performance que o diretor do filme queria. Entre estas especificações estão: acústica de ambiente, ruído de fundo e ângulos de visão e projeção.

**Taxa de bits** – Em telecomunicação digital é o número de bits que passa em um determinado meio de transmissão por certa quantidade de tempo (em geral, um segundo). A qualidade dos arquivos em MP3 é medida pela taxa de bits. 128 Kbps é considerada como a qualidade-base semelhante ao CD de áudio.

**Teorema de Nyquist** – Resumidamente, este teorema diz que: “a frequência amostral de um sistema de captura de áudio digital deve ser, pelo menos, o dobro da frequência mais alta contida no som analógico fonte”, caso contrário, *aliasing* irá ocorrer no som digital resultante. Este teorema foi desenvolvido no Bell Labs por C. Shannon e H. Nyquist. É daí que se originou o padrão de frequência amostral do CD de áudio comum (44,1 KHZ), que é cerca do dobro das frequências mais altas ouvidas pelos seres humanos.

**Tracker** – É como são chamados os *softwares* que compõem músicas ou sons no formato MOD.

**Upload** – Termo em inglês, em geral, tão ou mais utilizado do que aquele em português (“carregar”). É o ato de carregar um arquivo do computador de um usuário para algum local de uma rede de computadores.

**VBR** – Sigla para “Variable Bit Rate”. É o tipo de extração digital do som (popularmente conhecido por “ripagem”) em que a taxa de bits é ajustada conforme a complexidade do sinal fonte. Por exemplo, se há silêncio na música, a taxa de bits diminuirá e quando houver um complexo som sinfônico, a taxa de bits deverá aumentar, ou seja, os bits são alocados para os pontos da música onde se fazem mais necessários. A vantagem da VBR sobre a taxa constante de bits (CBR) é que ela oferece melhor qualidade de som com tamanho menor de arquivo.

**WMA** – Contração de Windows Media Audio, formato de som padrão da empresa Microsoft.

**WEB** – Contração de *World Wide Web* (Grande Rede Mundial). Tecnicamente a WEB seria a parte da Internet que tem capacidade multimídia, pois a Internet por si só já existia antes mesmo da estruturação da WEB ou WWW, quando era uma rede restrita a instituições governamentais, militares e universitárias na forma texto. Atualmente, WEB tem um sentido mais amplo e também é sinônimo de Internet.



**ANEXO- Lista de Arquivos de Áudio e suas Extensões conforme “The Sonic Spot” (THE SONIC SPOT, 2002) (o texto de descrição está em Inglês para melhor correlação com as extensões):**

<b>Extensão</b>	<b>Descrição</b>
669	Composer 669 module
669	UNIS Composer module
AIFC	Compressed Audio Interchange Format File
AIFF	Audio Interchange Format File
AIS	Velvet Studio Instrument
AKP	Akai S5000/S6000 Program File
ALAW	Raw A-law data
AMS	Extreme Tracker Module
AMS	Velvet Studio Module
APEX	AVM Sample Studio bank
ASE	Velvet Studio Sample
ASF	Microsoft Advanced Streaming Format
ASX	Microsoft Advanced Streaming Format Metafile
AU	Sun/Next Audio File (linear m-law or A-law)
AVI	Microsoft Audio Video Interleave File
AVR	Audio Visual Research sound file
C01	Typhoon wave file
CDA	CD Audio Track
CDR	Raw Audio-CD data
CMF	Creative Labs Music File
DCM	DCM Module
DEWF	Macintosh SoundCap/SoundEdit recorder instrument
DF2	Defractor 2 Extended Instrument
DFC	Defractor Instrument
DIG	Digilink format
DIG	Sound Designer I audio
DLS	Downloadable Sounds
DMF	Delusion Digital Music File
DSF	Delusion Digital Sound File
DSM	Digital Sound module

DSP	Dynamic Studio Professional module
DTM	DigiTrekker module
DWD	DiamondWare Digitized audio
EDA	Ensoniq ASR disk image
EDE	Ensoniq EPS disk image
EDK	Ensoniq KT disk image
EDQ	Ensoniq SQ1/SQ2/KS32 disk image
EDS	Ensoniq SQ80 disk image
EDV	Ensoniq VFX-SD disk image
EFA	Ensoniq ASR file
EFE	Ensoniq EPS family instrument
EFK	Ensoniq KT file
EFQ	Ensoniq SQ1/SQ2/KS32 file
EFS	Ensoniq SQ80 file
EFV	Ensoniq VFX-SD file
EMB	Everest embedded bank file
EMD	ABT Extended module
ESPS	ESPS audio file
EUI	Ensoniq EPS family compacted disk image
F32	Raw 32-bit IEEE floating point waveform values
F64	Raw 64-bit IEEE floating point waveform values
F2R	Farandoye Linear module
F3R	Farandoye Blocked Linear module
FAR	Farandoye Composer module
FFF	Gravis UltraSound PnP bank
FSM	Farandoye Composer WaveSample
FZB	Casio FZ-1 Bank dump
FZF	Casio FZ-1 Full dump
FZV	Casio FZ-1 Voice dump
G721	Raw CCITT G.721 4bit ADPCM format data
G723	Raw CCITT G.723 or 5bit ADPCM format data
G726	Raw CCITT G.726 2, 3, 4 or 5bit ADPCM format data
GIG	GigaSampler file
GKH	Ensoniq EPS (VFX, SD, EPS, ASR, TS) family disk image
GSM	Raw GSM 6.10 audio stream or raw 'byte aligned' GSM 6.10 audio stream
GSM	US Robotics voice modems GSM QuickLink/VoiceGuide/RapidComm
IFF	Interchange Format File
INI	Gravis UltraSound bank setup extract plus patch files

INS	Ensoniq instrument
INS	Sample Cell/II instrument
IT	Impulse Tracker module
ITI	Impulse Tracker instrument
ITS	Impulse Tracker sample
K25	Kurzweil K2500 (identicle to KRZ)
K26	Kurzweil K2600 (identicle to KRZ)
KMP	Korg Trinity KeyMap
KRZ	Kurzweil K2000
KSC	Korg Trinity Script
KSF	Korg Trinity Sample File
MAT	Matlab variables binary
MED	MED/OctaMED module
MID	Standard MIDI song/track information
MOD	Amiga SoundTracker / Protracker / NoiseTracker / Fastracker / Startrekker / TakeTracker module
MPEG MP2 MP3	MPEG-1 (Moving Picture Experts Group) Audio Layer I, II and III compressed audio
MT2	MadTracker 2 module
MTE	MadTracker 2 Envelope
MTI	MadTracker 2 instrument
MTM	MultiTracker module
MTP	MadTracker 2 Pattern
MTS	MadTracker 2 Sample
MTX	MadTracker 2 Extension
MWS	MWave DSP synth's instrument extract
NST	NoiseTracker Module
OKT	Oktalizer module
PAC	SBStudio II Package or Song
PAT	Advanced Gravis Ultrasound / Forte tech .patch
PBF	Turtle Beach Pinnacle Bank File
PRG	Akai MPC2000 Program File, WAVmaker program
PHY	PhyMod Physical Modeling data
PSM	Protracker Studio module
PTM	PolyTracker module
RA	RealNetworks RealAudio compressed streaming data
RAM	RealNetworks RealAudio Metafile

RAW	PCM signed raw audio
RBS	Propellerhead's Rebirth Song File
RMF	Beatnik's multimedia Rich Music Format
ROL	Adlib Synthesized Instrument Music file
RTI	RealTracker instrument
RTM	RealTracker module
RTS	RealTracker sample
S3I	Scream Tracker v3 instrument
S3M	Scream Tracker v3 module
SAM	MODEDIT sample file
SB	Raw signed PCM 8bit data
SBK	Emu Systems SoundFont Bank patch collection
SBI	SoundBlaster Instrument
SD	Sound Designer I audio
SD2	Sound Designer II flattened audio or data fork
SDK	Roland S-550/S-50/W-30 Disk Image
SDS	MIDI Sample Dump Standard
SDX	Sample Dump Exchange
SF	IRCAM Sound File"
SF2	Emu Systems SoundFont v2.0 patch collection
SMP	SampleVision audio, AdLib Gold Sample
SND	Akai MPC-series sample, PCM unsigned raw audio, NeXT Sound, Macintosh Sound Resource
SOU	SBStudio II audio
SPPACK	SPPack sound sample
STM	Scream Tracker Module 1 & 2
STX	Scream Tracker Module
SW	Raw signed PCM 16bit data
SYX	Raw MIDI System Exclusive message(s)
SYH	Synchomatic Instrument
SYW	Yamaha SY-85/SY-99 Wave audio
TD0	Akai Teledisk Sound Library
TXT	ASCII text parameter description
TXT	ASCII text formatted audio data
TXW	Yamaha TX-16W Wave audio
UB	Raw unsigned PCM 8bit data
ULAW	Raw m-law (CCITT G.711) data
ULT	UltraTracker module

UNI	UNIMOD module
UW	Raw unsigned PCM 16bit data
UWF	UltraTracker WaveSample
VOC	Creative Labs audio
VMD	Convex Raw sample
VMF	Convex SpeechThing / Voice Master sample
VOX	Dialogic ADPCM audio
W01	Yamaha TX16W or SY-series wave
WAV	Microsoft Windows RIFF WAVE
WFB	Turtle Beach WaveFront bank
WFD	Turtle Beach WaveFront drumkit
WFP	Turtle Beach WaveFront program
WOW	Grave Composer module
XI	Fastracker 2.0 instrument
XM	Fastracker 2.0 module
XP	Fastracker 2.0 pattern
XT	Fastracker 2.0 track